CRANFIELD UNIVERSITY

REABETSWE KGOROEADIRA

THE IMPACT OF COMMERCIAL PEER-TO-PEER LENDING WEBSITES ON THE FINANCE OF SMALL BUSINESS VENTURES

CRANFIELD SCHOOL OF MANAGEMENT

PhD Thesis

Academic Year: 2011 – 2014

Supervisor: Professor Andrew Burke

February 2014

CRANFIELD UNIVERSITY

CRANFIELD SCHOOL OF MANAGEMENT

PhD Thesis

Academic Year 2011- 2014

REABETSWE KGOROEADIRA

The impact of commercial Peer-to-Peer lending websites on the finance of small business ventures

Supervisor: Professor Andrew Burke

February 2014

This thesis is submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Abstract

In this dissertation, we set out to examine empirically the impact of commercial Peer to Peer (P2P) lending on the finance of small business ventures. Since, this is the first study that looks at the funding of small business ventures by commercial P2P lending website; we have collected and created a new and unique data set taken from Prosper.com, one of the dominating P2P lending websites; which formed the basis of our analysis. These data offer a unique opportunity to test theory - looking at information asymmetry problems and the mechanisms adopted to deal with them within a new context. The thesis comprises of three empirical chapters; we follow entrepreneurial finance literature in raising some of the key questions concerned mainly with: credit extension, the cost of credit and modeling default. We use robust analysis methods specifically: Probit, Tobit and 2-stage Heckman models to check for factors that drive credit allocation, factors driving the cost of credit and determinants driving default for small business loans.

General insights from our first empirical study shows that P2P lending depicts a new small business venture loan market, where previously underserved early stage entrepreneurs and those looking for small amounts are able to access unsecured credit through the relaxation of collateral. Although collateral is not required, we find that the supply of loans tends to flow to the least risky entrepreneurs; those who are homeowners, with high credit ratings. In our findings, we also demonstrate that firm level characteristics have little impact on loan supply while reducing information asymmetries through giving volunteering information improves access to loans. In general, our findings are both interesting and important as they suggest that P2P lending is a low risk form of debt finance. In this sense, lenders act like traditional debt financiers. However, the way in which they appraise funding opportunities characterise typical decision making of equity investors such as Business Angels and VC, who tend to focus more on people, rather than the business itself.

Findings from the second empirical study suggest that at an average lending of between 18 percent and 20 percent; P2P lending is a very expensive form of debt finance. Banks typically refuse to extend credit given such high interest rates as this tends to alter the borrower pool such that only the riskiest of borrowers have projects that generate returns that are high enough to be

able to re-pay these interest rates. In effect, the bank supply curve is backward bending above 10 percent on conventional terms of lending. Consequently, if we were to characterise P2P lending we would effectively conclude that it is typically a high cost finance with required returns expected to be likely in the levels of Business Angels and VC equity investments.

Finally, In terms of lender return and default, we find that the expected return to lenders is 3.26 percent, which is above the opportunity cost of capital in the US. Therefore, P2P lending is profitable from the investor point of view, albeit in a narrow sense. In general, the results suggest that average lenders on P2P platforms are amateurs, who actually have a higher risk tolerance. For these lenders, the risk of losing a small proportion (as little as $25) per investment in the overall portfolio of loans is offset by the potential gain from high interest rates charged for loans. Interestingly, our results show that return from the top 5 percent of lenders average at 6.1 percent per annum. Given the fact that P2P lending is generally a young market, and the fact that majority of lenders attracted to P2P lending are relatively uninformed amateurs in making investment decisions, the results suggest that if the amateur lenders do indeed learn, it then becomes plausible that in time the returns in this market may generally converge to be better (and gravitate towards the 6.1 percent achieved by top 5 percent). However, if the P2P lending platforms continue to attract a pool of amateur lenders, the average returns of 3.26 percent may render the market somewhat unsustainable in the long run.

Overall, our findings are novel, namely that P2P lending depicts a new venture loan market where previously underserved early stage ventures and those looking for small amounts are able to access credit; with the relaxation of typical collateral requirements. The big lesson however about P2P lending as a form of small business finance is that it really comes down to personal features rather than business features. Put another way, our findings suggests that the decision to extend credit and the pricing of loans in this context may possibly be relatively idiosyncratic - depending more on personal reputation of the small business owner than on the observed characteristics of the firm.

Table of Contents

List of Tables

List of Figures

Acknowledgements

First, I would like to thank my supervisor Prof. Andrew Burke for his guidance, support and continued encouragement throughout the PhD process.

I am also grateful to my PhD review committee, Dr. Stephanie Hussels, Prof. Mark Jenkins, Dr. Ronny Mason and Dr. Palie Smart for all their valuable comments and motivations at my first, second and third reviews. I am also thankful to my fellow PhD colleagues and all the staff of Cranfield School of Management for their continued support throughout this journey.

I am eternally grateful to my super hero mother, Blossom Kgoroeadira, for her unprecedented support, consistency, inspiration and love. The journey would never have been so fruitful without her being my 'cheer leader'. I would also like to thank my sister Mmoko and brother Leepo for always believing in me. To my life partner Busani, thank you for your patience. Finally, I offer my sincere gratitude to Cranfield School of Management for funding the study of my PhD degree.

Chapter 1 Introduction

1.0 Introduction

To successfully launch and grow a small business venture, business owners require adequate funding. The economics and finance literature provide strong evidence that sufficient starting capital is a binding constraint for potential small business owners. Entry into self-employment increases with a sudden increase in personal wealth, for example through inheritance (Holtz-Eakin et al, 1994; Blanchflower and Oswald, 1998; Burke, Fitzroy and Nolan, 2000). Likewise, absences of funds limit individuals to start businesses (Evans and Jovanovic, 1989). Evans and Jovanovic (1989), using the National Longitudinal survey from the United States, estimate that new business owners are limited by 1.5 times the size of their initial assets in starting a new business. Finance is thus a crucial element to new business entry.

The development and application of advanced information technology is altering the finance market (Han and Greene, 2007). Over the past few years, commercial peer-to-peer (P2P) lending websites have become a new innovative approach to mobilise and disseminate small business capital. Individuals, wishing to gain interest from their investment, extend credit to small business ventures directly, through the internet, without intermediation of traditional financial institutions like banks. Loan level data provided by Prosper.com (henceforth Prosper) and LendingClub.com (henceforth LendingClub) – the largest P2P lending sites – indicate that the dollar volume of P2P lending grew by nearly 300 percent between 2008 and 2011 (Figure 1-1).

...........Figure 1-1 goes around here...........

The total dollar amount that went to small business ventures has also seen an increase. Since 2008, LendingClub and Prosper have been responsible for over $100 million in small business loans (Table 1-1). Although the total dollar volume of P2P lending market is small relative to traditional sources of small business finance[1], this market is growing quickly and may represent an important niche in areas of external sources of small businesses funds. For Prosper, business

[1] According to the US Small Business Association, $588 billion in small business loans was outstanding as at 30 June 2012

loans represent 16.1 percent of all dollars lent over the period 2008 – 2011 coming from 2 million investors. For LendingClub, business loans are 5.6 percent of loan dollars, extended by just under 1 million individuals[2]. Prosper and Lending Club report that typical credit requests range from $1,000 to $35,000; while interest rates paid by borrowers normally range between 5 percent and 35percent.

...........Table 1-1 goes around here...........

Traditional lending institutions (banks, business angels, venture capitalists) face problems with regards to extending finance to small business ventures. Literature explains these problems largely to be accounted for by information asymmetries in capital markets, where borrowers are assumed to have more information about their prospective projects than lenders (Stiglitz and Weiss, 1981; Ang, 1991; 1992; Avery et al, 1998). If lenders are unable to determine the quality of the business venture because they lack full information, they raise average price of capital (interest rates in the case of banks). Because of the average high interest rates offered, low risk borrowers (knowing their worth) lack the incentive to access finance; they may opt to go look elsewhere. Stiglitz and Weiss (1981) argue that banks may therefore find it optimal not to raise interest rates in conditions of access demand because by so doing, they will worsen the quality of the borrower pool, a phenomenon known as adverse selection. This arises because only high risk borrowers can pay higher interest rates.

The change in interest rates may also influence borrowers to undertake riskier project, a phenomenon known as moral hazard, given that if the initiative becomes successful 'the business owner takes all'. Banks do not share in the profits or success of the undertaken project. If the initiative fails however, it is the banks that lose the funds they have extended to the business owners. Consequently, banks may opt to ration credit instead as an alternative of charging high interest rates. Since business start-ups and young small firms are the most likely to be

[2] www.prosper.com and www.lendingclub.com

'unknowns' to the banks, the problems of adverse selection and moral hazard would seem to be particularly acute for this cohort of business ventures.

Information asymmetries may also be problematic because of the relatively high fixed costs of gathering information lenders may incur for small transactions (especially in the case of venture capitalists and business angels); consequently lenders may opt to not extend credit to businesses looking for small amounts of capital.

To date, much of the research has focused on two main approaches by which traditional lending institutions attempt to cope with challenges caused by information asymmetries when extending finance to small firms: *signalling* and *relying on social ties*. The *signalling* approach emphasises the facilitative role played by:

- Collateral to distinguish between borrower types and to mitigate the risk of moral hazard when borrowers do not put enough effort into the business (Bester, 1985; Besanko and Thakor, 1987; Bester, 1987);

- Close relationships established between lenders and borrowers, which serve to improve information flow used to appraise credit risk (Sharpe, 1990; Petersen and Rajan, 1994; Berger and Udell, 1995; Cole, 1998);

- Human capital of the business owner; typically proxied by the owner's education level and work experience; in influencing the future prospects or performance of the business (Bates, 1991; Cressy, 1996).

In general, it is expect that individuals with access to collateral, with pre-existing relationships with lenders, who have greater work experience, education and knowledge of the market, signal better credit quality and hence are likely to access credit from lenders.

The *social ties* approach emphasises the facilitative role played by the small business owner's direct and indirect connections to potential capital providers (e.g. Hall and Hofer, 1993; Steier and Greenwood, 1995) and demonstrates that endorsements and social alliances with prominent third parties, which serve as a reputation gesture, can assist small firms in gaining access to finance (Stuart, Hoang and Hybels, 1999). Venture capitalists look for these endorsements when making funding decisions (Baum and Silverman, 2004); while microfinance institutions look to social ties to be able to implement joint liability lending (Hartley, 2010).

Despite mechanisms that may help overcome information asymmetries, theoretical arguments suggest that there are, nonetheless, a number of categories of small business ventures that could still be affected by information issues (Deakins *et al*, 2008; Coleman, 2000). For example business start-ups and young firms, because of their 'newness', may have little or no access to the aforementioned mechanisms available to indicate their credibility; yielding conditions under which serious information asymmetries prevail. Hence a funding gap develops. This realisation has prompted various forms of public policy interventions such as the loan guarantee schemes aimed to increase the supply of finance to small businesses. It remains unclear however if public policy can solve this funding gap issues (Storey, 1994). Hence, an innovation like P2P lending is of interest.

1.1 How small business start-ups are typically financed

The consequence of the aforementioned problems faced by traditional lenders in extending credit to small business ventures, especially firms in the early stages, is summarised in a framework put forward by Berger and Udell (1998) as illustrated by in Figure 1-2. Berger and Udell (1998) assert that in a typical distribution of business start-ups (in the US and elsewhere), for a significant proportion of firms, business capital comes from the 4Fs: first the Founders themselves reach deep into their own pockets for initial funds; next they turn to Family, Friends, and Foolhardy investors (also known as business angels).

...........*Figure 1-2 goes around here*...........

Bygrave and Quill (2007) confirm this trend through a study that was carried out across 42 countries. They find that 62 percent of the start-up finance came from the Founder(s) of the business; who self-fund through savings, second mortgage, credit cards etc. The remaining 38 percent of start-up funds came from external sources. For the proportion of business owners that receive start-up funds from external sources, the distribution is such that they predominantly obtain (equity) funds from family and friends. According to Bygrave *et al* (2003), 33 percent of the 'fastest growing' Inc. (500) companies in the US reported to have raised start-up funds from family and friends in the year 2000. Preston (2007) indicates that the amount that friends and family invest is modest, usually defined as less than $25,000. This form of finance typically represents the very first funds needed to finish the business plan, create a prototype, or conduct validating research.

A small proportion of business start-ups are able to raise finance from business angels - individuals with disposable wealth, who typically invest their own money into strangers' businesses. When business angels invest in early stage firms however, they typically target very high rates of return - as high as 25 percent to 47 percent (Lambert, 2010); which may preclude firms that do not fit these preferences of business angels. Some scholars (see for example Sohl, 1999) say that firms need to have the potential to create at least $10 million in sales after five years to be appropriate for angel investment. Consequently, business angels concerned with maximising returns, have sought to redistribute funds away from early stage ventures towards growth stages as shown in Figure 1-2 (Amatucci and Sohl, 2007). In terms of the size of investment, business angels typically provide funds in the range $25 000 to $500 000 (Shane, 2008; Benjamin and Margulis, 2000).

An even smaller proportion of small business ventures will raise start-up funds from primarily banks. Most banks tend not to make loans to new or young businesses because of the lack of assets, operational track record and other factors of inherent risk. But as shown in Figure 1-2, when the start-up firms grow, gain further experience and become less information opaque– banks are able to make lending decisions; hence they become more able to extend credit to these firms.

Lastly, only a miniscule number of small business ventures will ever tap into venture capital as a source of funds for their start-up business. Bygrave and Reynolds (2004) stipulate that as little as 1 in 10,000 small firms are able to attract venture capital in general; this dynamic is even more exaggerated for business start-ups and young firms (than in the whole population of small businesses ventures). Venture capitalists are more attracted to arguably even higher rates of returns than those expected by business angels. Given that the majority of small business ventures are not managed to pursue a growth strategy, which precludes the type of high-growth prospects that are so attractive to venture capitalists (Berger and Udell, 1998); these lenders tend to avoid business start-ups and young firms. Venture capitalists typically invest in the regions of over $500,000 with some deals in the range well over $10 million (GEM, 2004: 17).

Therefore for business owners looking for capital to start a business, it is likely that they will have to put up a large percentage of the funds from their own pockets. For those that look to external investors, informal investors in the form of friends and family seem to be the best bet; and if their businesses survives and reach the (post seed and) growth stage – informal investors in the form of business angels would be next in line to approach for funding. Then possibly after a while, when the start-up has built credit history and the strength of their books, banks may become a plausible source of funds. Eventually, if the start-up reaches growth phase, then it may or may not get venture capital.

Berger and Udell (1998) eloquently summarize the above argument as depicted in Figure 2-1. They show that small businesses can be placed on a size/age/ information continuum. Smaller and younger firms – more likely to face information asymmetry problems - lie near the left end of the continuum, indicating that they must rely on owner finance, friends and family, and/or angel finance (collectively known as informal equity).

1.2 Research motivation

Since Berger and Udell (1998) proposed their framework however, alternative small business funding mechanisms such as P2P lending have since entered the financial landscape. P2P lending is an alternative credit market that allows individual borrowers and lenders (who are strangers) to engage in credit transactions without going through traditional financial institutions like banks. Mostly, transactions take place through websites (known as platforms). These P2P websites provide capital for a variety of financial needs, including the provision of capital to small business owners. It is typically a form of unsecured debt finance. Unsecured debt finance is not typical at seed/start-up stage for small business ventures; finance at this stage is dominated by (informal equity). Hence, P2P lending is an innovation that may possibly open up the option of external debt finance to early stage ventures; especially if this market continues to grow.

In Figure 1-3, we show the typical typology of P2P lending. To obtain a loan, small business owners create a profile on the P2P lending websites, providing a small amount of information about their business. On most websites, the borrower stipulates (to potential lenders) the amount of money they wish to borrow and the interest rate they wish to pay for the loan. On some websites, the lenders make those determinations. Loan size can vary between $1,000 and $35,000; while interest rates can vary between 5 percent and 36 percent. On all websites, potential lenders decide how much, if any, funding to offer. In the case where lenders decide to offer funds, generally they spread their risk by backing a portion of each funded loan.

...........Figure 1-3 goes around here...........

The P2P lending landscape differs in at least three ways when compared to traditional lending institutions like business angels, banks, and venture capital. First, on P2P lending websites individuals extend loans to small business ventures without ever physically meeting the business owners; all transactions take place online. In traditional lending, however, the role of both physical contact and site visits (if need be) forms an important and integral aspect in credit extension; decreasing information asymmetries. It may be argued therefore that information asymmetries in the particular case of P2P lending might logically be expected to be even more severe.

Second, hundreds of potential lenders assess and screen the credit requests from small business venture at any given time. It could be argued therefore that individuals understand each other’s business better than traditional lenders. For example, some individuals may have knowledge about the area where the borrower wishes to start the business, others might have expertise in the product or the technology or the feasibility of the business. Hence, information asymmetries may potentially be less severe on P2P websites.

Third, traditional lenders like bankers are trained experts in conducting due diligence and assessing credit risk. In contrast, lenders extending credit to small businesses on P2P websites, although they have basic financial knowledge, it is questionable whether they have the sophistication and training to conduct efficient due diligence. Hence one might argue therefore that lenders on P2P websites may possibly be less informed in appraising credit risk and making investment decisions. It is plausible that they may be amateurs in appraising credit risk and making investment decisions.

Given these differences, ***are the aforementioned problems faced by traditional lending institutions in allocating finance to small businesses better or worse in the P2P lending context?*** This question is topical because it has yet to be established at an empirical level whether P2P lending merely crowds out traditional small business finance or whether it actually provides finance to businesses that otherwise would not get it (i.e. business start-ups and young firms; or businesses looking for small amounts of capital – unable to raise funds from business angels and venture capitalists).

1.3 Main aim of the dissertation, research questions and summary of findings

In this dissertation we explore the potential viability of P2P lending in the financing of small business ventures. Despite the continued growth of this phenomenon (i.e. millions of dollars from millions of individual lenders being extended to small business ventures), even basic academic knowledge of the dynamics of P2P lending in the financing of small businesses is lacking (outside of the still uncommon analysis of particular crowdfunding efforts see Burtch *et al*, 2011; Agrawal *et al*, 2011). For example, we know very little on the types of small business ventures this market serve, as well as the dynamics governing successfully raising funding in this context. We do not know whether P2P lending efforts reinforce or contradict existing theories about how small business ventures raise capital. There is also uncertainty about the long term sustainability of this market.

What's more, policy makers have also showed keen interest in P2P lending. For example, According to Solon (2012), the British government invested over $150 million in small business ventures through P2P lending websites in 2013, with the aim of "creating a more diverse financial infrastructure which better serves the needs of small and medium-sized companies" (Solon 2012:49). Similarly, the US government has recently passed a new legislation under the JOBS ACT in order to nurture and encourage growth of the P2P lending sector. The JOBS ACT allows start-ups and small businesses to seek funding of up to $1million per annum through P2P lending websites (The Economist, 2012). In short, this growing area of small business finance and government action is understudied; even as both practice and policy continue to rapidly advance.

1.3.1 Main research objectives

In this subsection, we outline the main research objectives. To put the main research objectives into perspective; we look to the area of entrepreneurial finance for some of the big research questions tackled (within the traditional 'offline' lending context). There are a finite number of big questions tackled in the entrepreneurial finance literature. There are questions concerned about which firms get funded and which firms don't; with the objective to look at where funding constraints take effect (Stiglitz and Weiss, 1981; Bester, 1985; Besanko and Thakor, 1987; Bester, 1987; Storey, 1994; Cressy, 1996; 2002; Parker, 2002; 2009; Cowling, 1997; Cowling et al, 2012; Mallick and Chakraborty, 2002). There are also issues or questions concerned about the cost that small ventures or early stage ventures pay for funding and what factors drive these costs (Sharpe, 1990; Petersen and Rajan, 1994; Berger and Udell, 1995; Cowling, 1997; Cole, 1998; Burke and Hanley, 2003; 2006,) and then there is literature around modeling default of small business ventures (Cowling and Mitchell, 2003; Berger *et al*, 2005; Agarwal *et al*, 2007; DeYoung *et al*, 2007; Berger and Gleisner 2009). In this dissertation we follow these three streams as a basis to set up our research agenda.

To the extent that the problems of information asymmetries and perhaps the effectiveness of mechanisms to counteract them might possibly differ in P2P lending, then the decision to allocate credit, the cost of credit and factors driving default may also differ. Due to the novelty of the phenomenon, currently there are no established mechanisms of coping with asymmetric information that have been identified among lenders financing small business on P2P websites. We argue therefore that if P2P lending responds to known signals established in traditional lending markets, it reinforces both the validity of these mechanisms in predicting borrower quality and the ability of 'amateur' lenders to select small businesses risk; which may render the market somewhat viable in extending small business loans.

Some of the mechanisms may not be available, for example P2P small business loans are unsecured – hence collateral in the form of Bester (1985, 1987) might not be available in this context. Also, loans are online – relationships with lenders as in the form of Petersen and Rajan (1994) are not available in this context. However, there are options available rendering new mechanisms of coping with asymmetric information that traditional lending institutions do not use. For example, large number of lenders which can aggregate the conducting of due diligence and lessen information asymmetries; as well as the wealth of personal and contextual information provided by the websites and the borrowers. These new mechanisms may be a source of advantage for P2P markets in reducing asymmetric information; hence rendering the market somewhat viable in disseminating credit to small business ventures.

Furthermore, in traditional markets many of the mechanisms available for existing firms are not available to new firms. This is a particularly important distinction as start-ups are widely perceived to be the riskiest class of small firms (Bates and Nucci, 1999; Evans, 1987; Dunne, Roberts and Samuelson 1989). Hence, we aim to distinguish whether mechanisms used to cope with information issues differ between new business start-ups and existing firms in the P2P lending context. We contend that if mechanisms used to cope with information asymmetries are no different for new business start-ups when compared to existing businesses in the P2P lending context, perhaps P2P lending may be introducing small business finance to a new market of small business ventures; previously under-served.

Overall, following entrepreneurial finance literature, one of the key objectives of the dissertation is to examine the determinants of credit allocation for small business loan on P2P lending websites. More specifically, in the first empirical study (Chapter 4) we address the *question (i) what factors drive the probability of funding for small business owners in the P2P lending context?* Since previous research has established that credit allocation to small business ventures is not only about extending funding, but also about the cost of funds (Petersen and Rajan, 1994; Cressy and Toivanen, 2001; Burke and Hanley, 2003; 2006); in the second empirical study (Chapter 5), we raise the question *(ii) what factors drive interest rates paid by small business owners in the P2P lending context?* Understanding both the issue of access to P2P funding and the cost of funding gives us the ability to ascertain whether P2P lending merely crowds out traditional small business finance or whether it actually provides finance to businesses that otherwise would not get it.

In addition to the credit allocation issues, another key objective of the dissertation is to shed light on the long term sustainability of the P2P lending market. The long term success of the P2P market primarily relies on the lenders' willingness to extend funds and the returns that they get for their effort. Rational, risk neutral and profit oriented lenders will only extend credit if they obtain at least as good a return in comparable alternative investments. Given the continued growth of P2P lending in general and an increase in the volume of small business loans in particular; a thorough analysis of loan default and loan profitability is needed so as to assess the future business potential of P2P lending websites. A question needs to be raised therefore of whether investing in P2P small business loans is beneficial for lenders. Hence, looking at the default activity in comparison to lender returns, (iii) *is it worth it for lenders to invest in small business ventures in this market?* We address this question in the third empirical study found in Chapter 6.

1.3.2 Summary of findings

The first empirical study titled: *Factors driving small business loan approval in P2P lending: who gets credit that doesn't* focuses on the determinants of access to credit in this new type of credit market. Specifically, we study data from Prosper.com. Because Prosper is the oldest and dominant P2P lending site, it is likely to serve as a broadly useful model for examining P2P lending efforts in financing small business ventures. The study is found in Chapter 4. To conduct the analysis, we use bivariate analysis and multivariate Probit regression models, based on a random sample of 12,526 loans requests issued between August 2007 and August 2013.

By way of preview, the bivariate results indicate that the typical firms approaching this market for small business funds are started or owned by borrowers who are less than premium credit risk[3]. These owners seem to be pursuing the business venture either as a side-line to their existing work or as a hobby given the fact that 60 percent of the sample indicate full time employment as their labour force activity. In terms of personal wealth, almost half of the business owners own their homes. New firms make up around 30 percent of businesses approaching this market; the remaining 70 percent is made up of already established businesses. Relative to a representative sample of US small firms of which new ventures form around 10 percent, new firms are over represented in the P2P context, suggesting that P2P lending may be catering to the funding gap experienced by early stage ventures. In terms of industry distribution,

[3] Over 60% of borrowers have had previous delinquencies and 1 in 3 have failed to meet previous loan obligations

82 percent of the sample is found in the retail, services, or finance industries. This seems typical of the sample of U.S. small businesses; as 87 percent of U.S. firms are represented in these sectors[4]. Finally, on average these firms are looking for small amounts of money ($10,430); they are willing to pay a high price for credit (24 percent) with the majority (55 percent) looking for working capital.

The general insights from our first empirical study reveals that only 11 percent of loan requests manage to get funded in this market. Although collateral is not required, we find that the supply of loans tends to flow to the least risky entrepreneurs; those who are homeowners, with high credit ratings. Hence, reputation is the single most important determinant of loan supply. In our findings, we also demonstrate that firm level characteristics (including age of the firm) have little impact on loan supply. Overall, these findings are both interesting and important as they suggest that P2P lending depicts a new small business venture loan market, where previously underserved early stage entrepreneurs and those looking for small amounts are able to access unsecured credit through the relaxation of collateral. Our results also suggest that P2P lending is a low risk form of debt finance. In this sense, lenders act like traditional debt financiers. However, the way in which they appraise funding opportunities characterise typical decision making of equity investors such as Business Angels and VC, who tend to focus more on people, rather than the business itself.

Interestingly, with regards to the mechanisms unique to this context we introduce a new feature in the form of 'crowds' that do indeed help to reducing information asymmetries and adverse selection issues. Prospective lenders perceive loan requests attracting a large number of potential lenders to have conducted a great amount of due diligence. Furthermore, our findings suggest that the loan supply in this context may also possibly be relatively idiosyncratic, depending more on lender peculiarity like philanthropy. This observation is supported by evidence we found that the way P2P lenders generally appraise funding opportunities is more focused on people rather than the business itself. These findings are both interesting and important as they suggest new mechanisms of dealing with information asymmetries previously not considered by theory. We also find interesting results to the information in the pictures. All else equal, we estimate that

[4] Following traditional small business literature See ole,1998; we defined industry based on 1 digit SIC code

requests from borrowers that include a picture are 2 percentage points more likely to receive funding. Compared to the average probability of funding of 11 percent; this represent an almost 20 percent increase in the likelihood of receiving funding. Perhaps this emphasizes the importance of humanising the lending process on P2P websites.

Our results from the first study have key implications to theory. To recap, Stiglitz and Weiss put forward four key parameters underpinning their theory of information asymmetry and adverse selection: collateral, conducting due diligence, refraining from high interest rates to avoid moral hazard and adverse selection issues, and the inferred face to face borrower-lender interactions. From our results, we find that some weights of these parameter values are likely to change in Stiglitz-Weiss model.

First, the general insight we get from our first study is that borrower reputation, stipulated by credit grades, is the single most important determinant of credit allocation .The significance of using the credit grade helps to reduce the problem of adverse selection. The cost of defaulting will result in poorer scores – which are quantifiable to 24 percent in reduction of probability of funding. Moreover, with the advancement of internet, reputations which were previously limited within the 1-to-1 lending typology from banks, where if a borrower defaults on credit in one region or country for example, would not have an effect if the borrower were to move to another country has since have completely changed. But with internet age loan default may quickly go viral. The consequence of 1-to- many borrower-lender interactions relative to reputation over the internet and the ease with which default may go viral makes reputation to be a very important aspect within the P2P lending context.

Second, we see from our results that collateral, which was such an important determinant in reducing adverse selection issues in Stiglitz-Weiss theory, in the P2P lending it is unimportant.

Third, the general insight we get from our first study is that due diligence, although still an important factor, in the P2P lending context is conducted by the 'crowd'. This new feature, unique to P2P lending was not taken into consideration in the Stiglitz and Weiss framework – where credit risk appraisal was done by relatively one person. Consequently we introduce collecting intelligence as a means of eradicating information asymmetry and adverse selection issues. Furthermore, our results shift focus from 1-to-1 physical interactions between borrowers and lenders inferred in Stiglitz-Weiss theory and highlights the 1-to many borrower lender typology over the internet. Effectively rendering physical contact, which was previously seen an important aspect of reducing information asymmetry and adverse selection issues in theory, relatively less important.

Furthermore, In terms of due diligence, the crowd also introduces another distinctive change based in the notion that lenders in this context may have philanthropy ambitions which they may consider when appraising credit risk. For example, a lender taking into consideration philanthropic ambitions may look for different credit risk when compared to a lender whose sole ambition is to maximise returns. This may effectively alter access to credit and subsequently the P2P borrower pool.

In sum, we contend that the lessons we leant from our results about asymmetric information and adverse selection issues are quiet different to those developed in Stiglitz-Weiss model. Reputation is very important in this market (this was not highlighted by Stiglitz). We learn that physical contact and collateral becomes less important in reducing information asymmetries. We also learn about the importance of three key new features: collective intelligence of the crowd, an aspect of philanthropy present when appraising credit risk and the general element of fun which may drive lenders when choosing to allocate credit. We contend that Stiglitz-Weiss model will have to be updated to take into consideration the facts raised above in their theory of information asymmetry in order to reflect our finding.

Our results also have key practical implications. The fact that P2P lending relaxes collateral requirements and the fact that reputation is the single most important variant in P2P lending renders our results generalizable to the microfinance institutions, especially those in developing countries. The relaxation of collateral makes P2P a viable alternative to small business funding in developing countries, especially given the fact that it is in these regions where the wealth distributions tend to be skewed unfavourably. The importance of reputation simply highlights infrastructure that already exists, used by microfinance in developing countries. Hence this renders our results somewhat generalizable. The fact that the internet is an underlying layer in operationalizing P2P lending in developing countries, perhaps there is scope of using mobile smart phones instead of traditional computers – given that this technology is already entranced in developing countries.

The second empirical study titled: *Factors driving the cost of credit in P2P lending* is found in Chapter 5. The main purpose of the second study is to investigate the determinants driving the interest rate paid by small business loans on P2P websites. We adopt a Tobit model to assess the factors that drive interest rates paid based on a sample of 1417 funded loans from Prosper, issued between August 2007 and August 2013.

The general findings from the second empirical study suggest that at an average lending of between 18 percent and 20 percent; P2P lending is a very expensive form of debt finance. Banks typically refuse to extend credit given such high interest rates as this tends to alter the borrower pool such that only the riskiest of borrowers have projects that generate returns that are high enough to be able to re-pay these interest rates. In effect, the bank supply curve is backward bending above 10 percent on conventional terms of lending. Consequently, if we were to characterise P2P lending we would effectively conclude that it is typically a high cost finance with required returns expected to be likely in the levels of Business Angels and VC equity investments.

Further insight we get from our second study is that borrower reputation, stipulated by credit grades is the single most important determinant of the cost of credit .The significance of using the credit grade helps to reduce the problem of moral hazard. The cost of defaulting will result in poorer scores – quantifiable to an increase of 80 basis points. Second, we see from our results that collateral, which was such an important determinant in reducing moral hazard in Stiglitz-Weiss theory, in the P2P lending it is unimportant. The fact that P2P lending relaxes collateral requirements and the fact that reputation is the single most important variant in assessing the cost of credit on P2P websites, highlights some of the facts that were not previously considered by previous theory. That is, moral hazard may be solved by other means in the form of credit scores.

So, looking at the results of the first and second study collectively, if within the P2P lending market it is the relatively low risk borrowers who get the loans, then a natural question to address is why couldn't these low risk borrowers' access conventional bank loans? Moreover, why were they willing to offer high interest rates? One probable explanation is the condition of the credit market. Our analysis is based on data covering the period 2008 - 2013, arguably characterised as the height of the recession. According to SBA (2012) small business lending from banks constricted during this period. Suddenly, businesses who were perfectly good risk able to access credit from banks during buoyant economic times, were suddenly unable to tap into bank lending because banks tightened their lending standards to small firms. Subsequently, borrowers, keen to have access to capital, were hoping to attract possible lenders with their good credit record and high interest rates; just so they can keep their businesses afloat. Consequently, we find that P2P lending platforms provided an alternative form of venture finance in times of recession, when traditional financiers were unavailable.

The third and final empirical study titled: *Factors driving small business loan default: is it worth it worth it for lenders to invest* is found in Chapter 6. The objective of this study is to shed light on the long term sustainability of the P2P lending market. For P2P lenders, it is difficult to judge the quality of the deal offered beforehand, because lenders bear the default risk and they are not trained experts in risk management. Moreover, lenders on P2P websites tend to demand relatively high interest rates (average 18.5 percent) in order to compensate for credit risk. In the study, we address the raised issue by investigating the determinants of loan default for a cross section of 1417 small business loans.

Overall, the results show that on average 1 in every 4 loans funded in this market will result in default. Our analysis suggests that the most important determinant of whether borrowers repay their debts is the loan size they take up, the interest rates paid; their credit grade and whether they shared extra information upfront (in the form of pictures). In general our results also find that, contrary to the hypothesis that new business are a more riskier class, new firms are no more likely to default when compared to existing firms in this context. We also find that default does not vary across different types of industries such that P2P lending becomes sustainable only in some parts of the small business sector and not in others – all else equal; industry variables are insignificant determinants of default.

So, is it worth it for lenders to extend funds to small businesses in this market? In terms of lender return, we find that the expected return to lenders is 3.26 percent, which is above the opportunity cost of capital in the US. Therefore, P2P lending is profitable from the investor point of view, albeit in a narrow sense. In general, the results suggest that average lenders on P2P platforms are amateurs, who actually have a higher risk tolerance. For these lenders, the risk of losing a small proportion (as little as $25) per investment in the overall portfolio of loans is offset by the potential gain from high interest rates charged for loans. However, default is found to be related to risk as predicted by conventional theory, in both formal credit rating and also additional information and track records.

Moreover, a key finding is that default seems to also be associated with the cost of credit, which presents an interesting dilemma for lenders. Consequently, at a first glance, we inferred from our finding that lenders would have to consider a potential trade off between lowering the interest rates they charge borrowers in order to circumvent default. However, it is also plausible that some lenders extend credit to entrepreneurs in this context for idiosyncratic reasons – for example as some form of gambling. In this case, lenders would simply accept loan requests from borrowers offering high interest rates knowing that if the borrowers pay back the loan, they win big. However, if some of the loans in the portfolio result in default, the loss is not too big. Likewise, some lenders way choose to extend credit to entrepreneurs in this context driven by philanthropic desires, knowing that extending credit to the entrepreneurs would simply be helpful to their course.

Interestingly, our results show that return from the top 5 percent of lenders average at 6.1 percent per annum. Given the fact that P2P lending is generally a young market, and the fact that majority of lenders attracted to P2P lending are relatively uninformed amateurs in making investment decisions, the results suggest that if the amateur lenders do indeed learn, it then becomes plausible that in time the returns in this market may generally converge to be better (and gravitate towards the 6.1 percent achieved by top 5 percent). However, if the P2P lending platforms continue to attract a pool of amateur lenders, the average returns of 3.26 percent may render the market somewhat unsustainable in the long run. Given all that has been presented so far, P2P lending may come across as an inefficient use of resources. For example, if we were to follow the line of argument that lenders may choose to engage in P2P lending for idiosyncratic reasons such as fun or gambling etc and philanthropic reasons, it then becomes reasonable to infer that perhaps P2P lending may possibly result as an inefficient use of resources. However, the fact that P2P lending is profitable for an average, risk loving, uninformed lender, may suggest that this instrument may somewhat be an alternative form of financial instrument (albeit with some form of inefficient tendencies).

The third study makes key contributions. First, we shed light on the determinants of default in the P2P lending context; highlighting that borrower reputation continues to be the single most dominant determinant of default. Second, we find that the breed of investors extending credit to entrepreneurs in this context are amateurs, who have high risk tolerance, with a completely different utility function to that of lenders inferred in the model by Stiglitz and Weiss (1981). This study sheds some light that P2P lending may be availing a different type of investor to that previously seen in the small business lending literature. Finally, Caution will to be exercised when ascertaining whether our findings can be generalised to developing countries for example. As it stands, P2P lending does not have other mechanisms at its disposal such as peer pressure and a sense of community, which help keep default rates down when extending microfinance. Given that relationships in this context are very weak at best, since lending takes place online, default rates may therefore sky rocket in the context of Microfinance institutions, thus rendering P2P lending an unsustainable form of extending credit to small businesses.

1.4 Structure of the dissertation

The remainder of the dissertation is structured as follows: In Chapter 2, we review the literature, and set the context of the research; position the research questions and developed hypotheses. In Chapter 3 we discuss the data source and data set used in the analysis; then we put forward details of the methodology adopted in analysing the data. In Chapter 4 we present the first empirical study - concerned with the determinants driving the credit extension decision. Next in Chapter 5, we present the determinants driving the interest rates paid; whilst Chapter 6 presents factors driving default activity. Finally, in Chapter 7, offer final conclusions, contribution of the study and limitations of the study.

Table 1-1: Small business lending by Prosper and Lending Club, 2008 – 2012

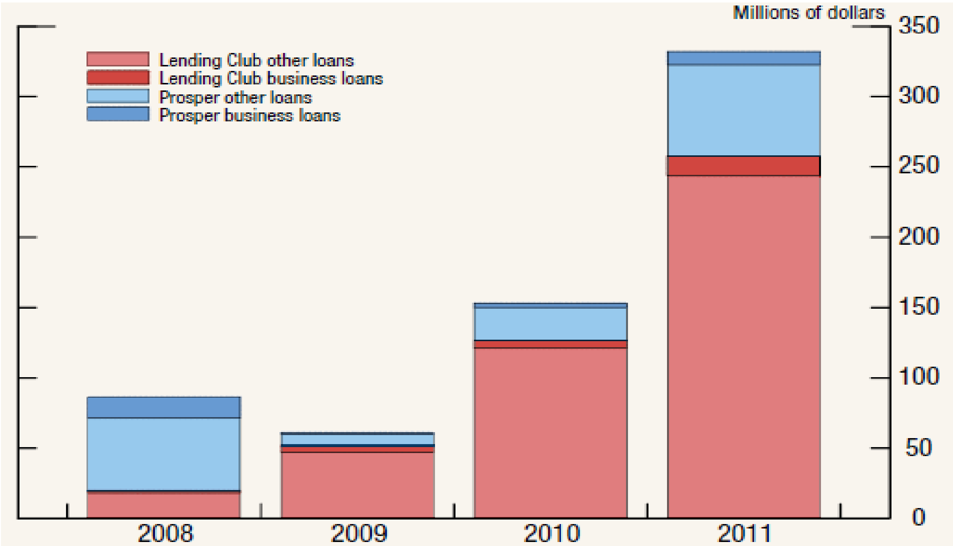
Number of loans (thousands) and dollar amount (millions) disseminated to small business ventures by Prosper.com and Lending Club, the two largest P2P websites for the period 2008 - 2012

Lender and Year	Number of loans	Dollar amount of loans
Prosper		
2008	1714	15,240,122
2009	212	1,165,140
2010	550	3,098,768
2011	1198	9,132,100
2012	1682	15,051,086
Lending Club		
2008	127	1,683,250
2009	358	4,392,125
2010	466	5,384,875
2011	975	13,861,950
2012	1386	22,502,666

Data source: www.prosper.com and www.lendingclub.com

Figure 1-1: Funded loans on Peer-to-Peer lending websites 2008-2011

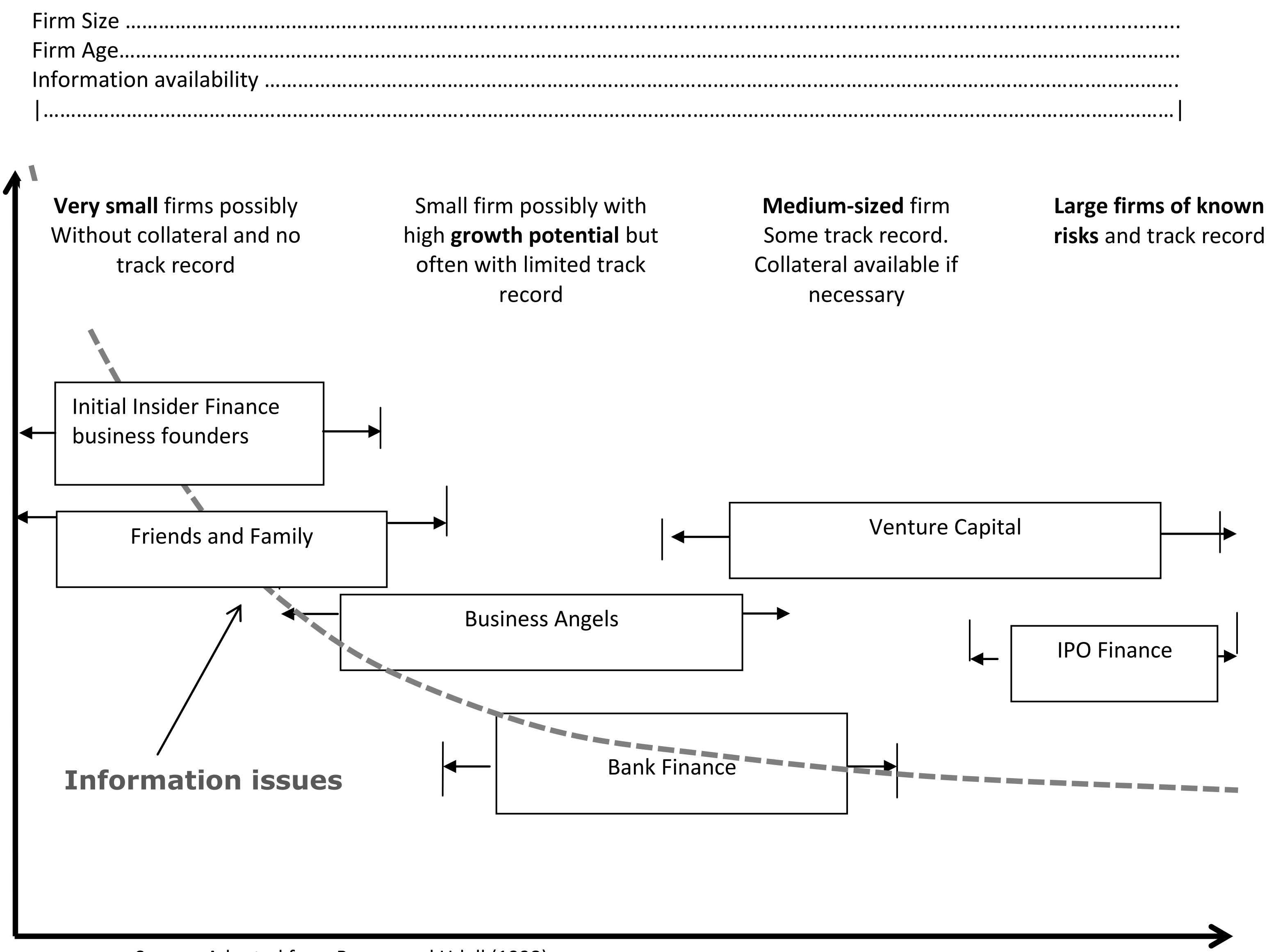
Small business loan composition (relative to general loans) for the period 2008 - 2011 on the two largest P2P lending websites, Prosper and Lending Club



Source: Board of Governors of the Federal Reserve System Report to congress on availability of credit to small businesses (2012: pp. 40)

Figure 1-2: Sources of small business finance

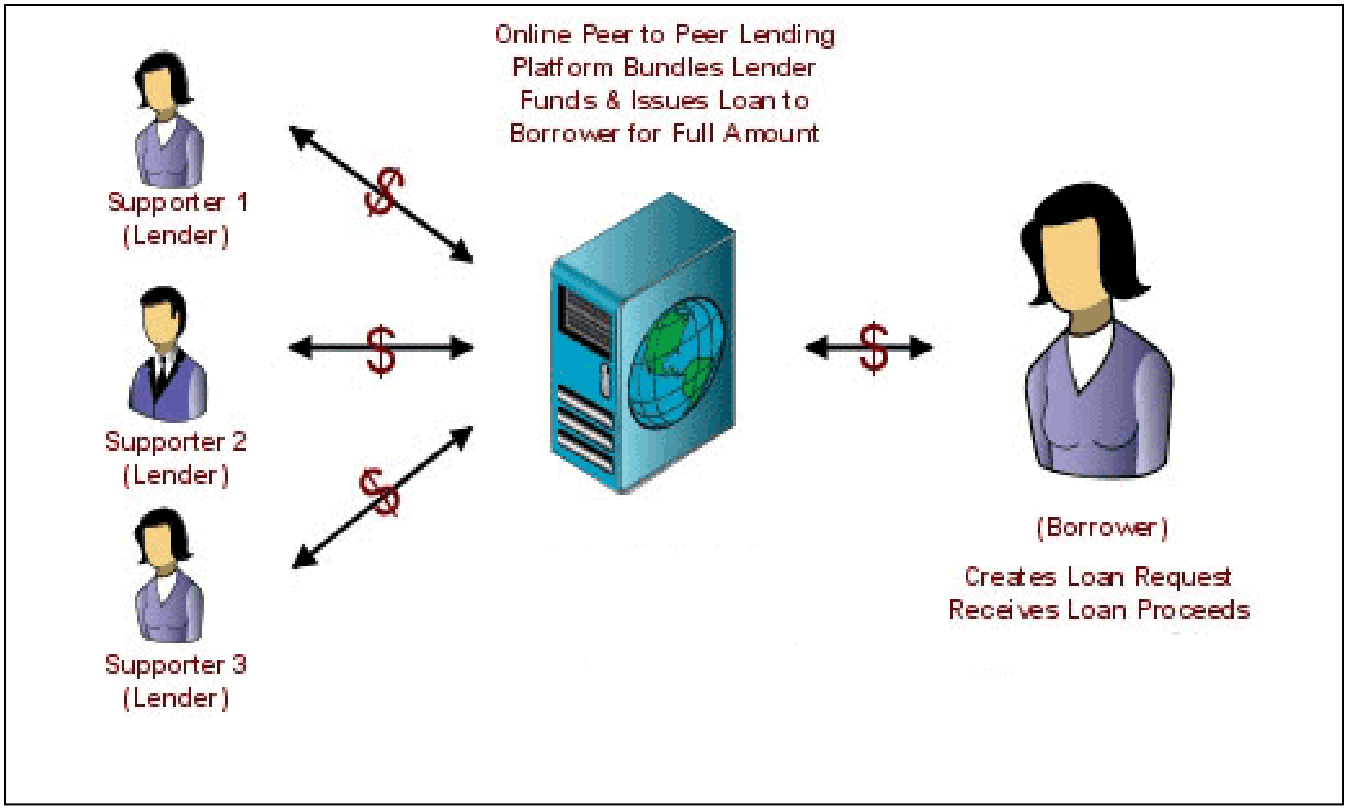
Different sources of finance for small business ventures by firm size, firm age and level of information availability across the business lifecycle



Source: Adopted from Berger and Udell (1998)

Figure 1-3: P2P Lending Landscape

Typical P2P lending landscape where the loan for a single borrower is characteristically funded by combining multiple lenders who contribute a small amount each without intermediation from traditional institutions like banks, in return they receive a pro-rata interest rate on their investment



Source: http://www.hacktrix.com/social-lending-websites-to-get-peer-2-peer-p2p-loans

Chapter 2 Literature

2.0 Introduction

This Chapter reviews literature on mechanisms typically used to mitigate the problems caused by asymmetric information in financing small business ventures. An extensive theoretical and empirical literature exists that considers mechanisms that may be used by prospective lenders to distinguish between borrower types and to mitigate the risk of moral hazard (when borrowers do not put enough effort into their businesses). These mechanisms include the provision of collateral that lenders can seize if borrower default (Bester, 1985, Besanko and Thakor, 1987; Bester, 1987); the development of close relationships between the lenders and the borrowers to improve information flow (Sharpe, 1990; Petersen and Rajan, 1994; Berger and Udell, 1995; Cole, 1998) and signals regarding the future prospects for the business based on borrower human capital (Cressy, 1996).

To put this literature into perspective, we draw from some of the big issues or questions raised in entrepreneurial finance literature. The literature is typically organised around 3 major issues as shown in Table 2-1: there are issues or questions concerned about which firms get funded and which firms don't; with the objective to look at where funding constraints take effect (section 2.1); there are issues or questions concerned about the pricing or cost of small firm finance - how much small ventures or new ventures pay for funding and what determines these costs (section 2.2); and then there is literature around default (section 2.3). In Table 2-1 the three fundamental issues are listed along with the smaller questions related to each stream. Each of these research streams is discussed and key papers are described highlighting mechanisms adopted to mitigate information issues.

...........Table 2-1 goes around here...........

The second part of the Chapter discusses P2P lending literature (section 2.4). Since research in this area is relatively new, the concept of P2P lending itself does not enjoy a unique definition in the literature. Therefore we start in section 2.4.1 by defining the concept of P2P lending, and then we identify the benefits and costs entailed by the lenders and the borrowing firms. Next, in section 2.4.2 we briefly review some of the literature concerned with P2P lending and small businesses finance. We conclude the Chapter by indicating the research gap which then sets the scene for raising our key research questions.

2.1 Credit extension and small business ventures

In this subsection we discuss mechanism typically used by lenders to attenuate information asymmetries when making credit extension decisions. A large literature exists that stipulates that personal wealth held by business owners may influence how a business is funded (Stiglitz and Weiss, 1981; Wette, 1983; Bester, 1985; Besanko and Thakor, 1987; Avery *et al*, 1998). Information on personal wealth (usually proxied by collateral) can improve underwriting decisions and lessen the extent of exposure for the lender. For example, with potential adverse selection where borrowers have superior information to that of the lender, creditworthy borrowers can use collateral as a signal of their quality (Bester, 1985). Likewise, if borrower effort creates moral hazard risk, then collateral may mitigate the incentives for unnecessary risk taking as borrowers realise that their personal wealth is at stake.

Theoretical models focusing on the signalling function of collateral follow from early contribution by Bester (1985) with modifications by Besanko and Thakor (1987). Given the variability in individual default risk, these studies show that borrowers with low probability of default (i.e. low risk borrowers) will reveal themselves by accepting collateral, which would be unattractive for high risk borrowers as collateral is costly. A similar argument holds in the case of moral hazard where collateral requirements serve as an incentive mechanism because higher collateral enforces borrowers to select less risky projects.

Some theoretical studies consider the effect of collateral in isolation (Stiglitz and Weiss, 1981; Wette, 1983); while contributions by Bester (1985, 1987) and Chan and Kanatas (1985) show that by treating collateral requirements together with variations in interest rates, collateral is negatively related to the borrower's risk. Hence, all else equal, borrowers with high probability of default prefer contracts with higher interest rates and lower collateral than borrowers with low default risk. The reason being that high risk borrowers can afford to pay higher interest rates. Moreover, they are also more likely to lose their collateral if their project fails.

The studies by Bester (1985, 1987) are based on the assumption that collateral is readily available to borrowers. Besanko and Thakor (1987) relax this assumption and show that credit rationing resurfaces when the borrowers face constraint on collateral availability. However, their model confirms that collateral nevertheless hold as a signalling devise as it does reduces rationing, even if it cannot eliminate rationing all together. In a companion paper, Besanko and Thakor (1987b) permit loan size to be used as a signal in conjunction with collateral and loan interest rates and show that the loan size is increasing in the borrower's (truthfully) revealed success probability.

The empirical evidence based on the above theories examining the association between borrower risk (broadly defined) and collateral is mixed. Studies by Berger and Udell (1990; 1995); Leeth and Scott (1989); and Boot *et al*, (1991) found that it is the riskier borrowers who are likely to be asked to provide collateral. Berger and Udell (1995), using interest rate premium as a measure of borrower risk, find that collateral is associated with higher risk premiums among small business loans. This result seems counter to the prediction of the theories put forward by Bester (1985) and others as argued above that high quality borrowers pledge collateral and opt to pay low interest rates. However, Berger *et al* (2011) put forward that it may be the case that collateral differences more often reflects observed quality differences, rather than unobserved differences between borrower types. Studies by Machauer and Weber (1998) report that collateral is independent of borrower type; while a recent study by Jiminez *et al* (2006) show that collateral is negatively related to ex post default on loans offered to young firms. The authors argue that ex post default may reflect high unobserved risk and hence ex ante information asymmetries.

Whatever the reason behind the mixed results and the lack of practicality of some of the models that describe the relationship between collateral and credit risk, there is sufficient empirical evidence to indicate the importance of collateral in credit markets with asymmetric information. Evans and Jovanovic (1989) show that absences of funds inhibit individuals to start businesses; similarly, studies by Holtz-Eakin *et al*, (1994); Blanchflower and Oswald, (1998); and Burke *et al*, (2000) show that entry into self-employment increases with a sudden increase in personal wealth.

Another mechanism available to reduce information problems when extending credit to small firm is through relationship lending (Boot and Thakor, 1994; Petersen and Rajan, 1994; Berger and Udell, 1995; Cole, 1998; Harhoff and Körting, 1998). According to this literature, lenders acquire information over time through contact with the firm, and/or its owner and use this information in their decision to extend credit. The premise is based on the fact that borrowers will be able to build a reputation over time where lenders are able to observe their repayment behavior. Hence, firms with a strong relationship with their prospective lenders are more likely to receive credit.

Traditionally, these studies measure the strength of the relationship in terms of its length, for example, the amount of time the bank has provided loan or other services to the firm (Petersen and Rajan, 1994, 1995; Scott and Dunkelberg, 1999). In general these studies report that strong ties with lenders lead to greater availability of credit for small firms (Petersen and Rajan, 1995; Berger and Udell, 1995; Harhoff and Körting, 1998). The studies by Cole (1998) and Machauer and Webber (2000) put an emphasis however that it is not only the strength of this relationship that is important, but the simple presence of the relationship between the lender and the small business that matters in the credit allocation decision.

Banking literature has demonstrated that observable characteristics of the small business and those of the business owner are also found to be highly valuable in reducing information problems. This is especially true for new business start-ups or small businesses at early stages of firm development when owners are usually the major (and probably the only) decision makers - suggesting that owner's patterns of financial services mimic those of the firm (Cassar, 2004). Borrower credit rating, education, age, and business experience are some of the most common attributes studied to evaluate the impact of the owner's characteristics on access to finance (Berger *et al*, 2005; Berger and Frame, 2007; Cressy, 1996; Burke *et al*, 2000). The general consensus in this literature is that educated borrowers, with high credit ratings and industry experience, who are much older (and therefore have longer track records and perhaps even possible savings), are likely to be extended credit. Likewise, Cole (1998), based on US data and Coleman (2010), based on UK banking data, report that firm age and firm size (measured in different ways) also influence credit allocation such that older firms and bigger small firms were more likely to be extended credit relative to new and smaller firms.

We sum up the discussion so far as shown in Table 2-2. The reviewed literature on lending under asymmetric information predicts that wealthier business owners (with access to collateral); with a good credit history (proxied by high credit ratings), who have previous existing relationship with potential lenders; who also have greater work experience, and are educated; generally signal better credit quality and hence are likely to access small business funds from lenders. Moreover, there is also evidence that suggests that firm age and size affects credit access such that older and larger small firms will have more access to credit.

...........Table 2-2 goes around here...........

Previous research has established that lending to small business ventures is not only about extending funding, but also about the cost of funds i.e. the interest rates paid (Petersen and Rajan, 1994; Cressy and Toivanen, 2001; Burke and Hanley, 2003; 2006). Therefore it remains relevant to gain an understanding of what drive the cost of credit for small firms. We now move on to discuss this literature in the next session.

2.2 The cost of small firm finance

Studies concerned with factors that influence interest rates paid by small business borrowers postulate that closer relationships with creditors improve information flows which may allow more accurate assessment of risk; and reduce information asymmetries which leads to lower rates of interest (Petersen and Rajan, 1994; Berger and Udell, 1995; Cowling, 1997; Harhoff and Körting, 1998; Keasey and Watson, 2000; Cressy and Toivanen, 2001).

Petersen and Rajan (1994) are among the first to examine empirically how bank-firm relationships affect the cost of funds using a sample of small privately held firms. The data comes from the 1988 National Survey of Small Business Finance conducted by the U.S. Small Business Administration and the Federal Reserve. They use three different measure of strength of relationships, namely: duration, the number of financial services (scope) and the number of lenders. They find a reduction of the interest rate among those enterprises that work with fewer institutions, although they didn't find a significant link between the duration and scope of the relationship and the price of debt. Berger and Udell (1995) use the same dataset as Petersen and Rajan (1994) but restrict the sample of loans to lines of credit. The reason is that lines of credit are more likely to be relationship loans than other types of loans. They find that borrowers with longer banking relationships pay lower interest rates (and are less likely to pledge collateral).

Studies by Keasey and Watson (2000) and by Cressy and Toivanen (2001), both based on data derived from a representative UK banks, also find that existence of a relationship with a creditor enables small firms to be charged lower interest rates. Contrary to the studies above, Harhoff and Körting (1998), based on survey data from 1509 German SMEs, find that interest rate is not significantly affected by duration of the relationship. The proxies of strength of relationship used are duration, the number of lenders and qualitative response in which firm managers indicate to what extent they consider their bank relationship as being characterised by mutual trust.

There is evidence that the interest rate charged to small firms incorporate whether the borrower provides collateral; such that borrowers who provide collateral are afforded lower interest rates. Burke and Hanley (2003, 2006) argue however that the collateral-interest rate relationship is not necessarily linear. Based on UK banking data and estimated with OLS regressions, they observe a U shape relationship between wealth and interest rates such that the cost of a loan becomes more expensive for borrowers whose personal wealth exceeds the collateral value and for those on the other extreme end of the spectrum who are less wealthy.

There is also evidence that interest rate charged by lenders incorporates firm specific characteristics such as firm industry, firm size and firm age (Keasey and Watson, 2000; Cressy and Toivanen 2001; Cowling, 1999). Keasey and Watson (2000) study the pricing of small firm loans based on UK bank data, and opt to include industry dummy variable so as to ascertain whether interest rates are driven by the industry the small business operates in. They find that none of the individual coefficients on the industry dummy variables influence the interest rate paid by small businesses. In terms of loan characteristics (i.e. loan size, loan purpose and loan term); studies by Cressy and Toivanen (2001) and Burke and Hanley (2006), all based on UK bank data, report a negative relationship between loan size and interest rates paid by entrepreneurs. Cressy and Toivanen (2001) however reports a reversed sign when endogeniety is considered in two stage equations. In terms of loan purpose, studies by Cressy and Toivanen (2001), Cowling and Mitchell (2003) and Burke and Hanley (2006) find that working capital loans exhibit higher default rates; hence they observe a positive significant relationship between working capital loans and interest rates paid.

Information about the cost of second round finance has also been presented in literature. A number of theoretical papers predict the relative cost of finance in the second borrowing period

(Petersen and Rajan, 1995; Boot and Thakor, 1994; Greenbaum *et al,* 1989; Sharpe, 1990; Diamond, 1989). Diamond's (1989) multi-period model postulates that a bank will reward survivors (those who chose lower risk projects or who were ex ante lower risk borrowers) by reducing their interest margins on second round finance. If entrepreneurs are aware of this incentive, then a bank should draw a safer pool of applicants for finance (reduced adverse selection) and borrowers will be persuaded to choose lower risk projects (reduced moral hazard). Boot and Thakor's (1994) model predictions concur with the description of cheaper second round finance.

However, other theoretical models predict that second round finance will be more expensive. Greenbaum *et al* (1989) and Sharpe (1990) outline models predicting rising costs as a firm's reputation becomes established. Sharpe describes an exploitative bank relationship where a bank exploits its information monopoly on a borrower, when information about a borrower's reputation cannot be observed by competing banks. This information monopoly is reflected in an interest margin hike for second-round finance. Sharpe (1990) argues that this interest margin hike is due to the entrepreneur being "informationally captured" by their lender.

Empirical studies by Binks and Ennew (1998) based on US survey data, show that longer relationships can lead to increased interest rate charges due to banks taking advantage of the firms' lock-in to the relationship. Similarly Hanley and Crook (2005), explicitly consider the impact of relationship on the cost of funds (while controlling for collateral) using a dataset for 1409 commercial loans in the year 1998 provided for a UK retail bank. They propose a two equation model for the joint determination of collateral and interest rates. They report higher interest rates for follow-up loans, that is, when there is an on-going relationship with the lender. They interpret this result, similar to Sharpe (1990), as evidence of a lock in effect.

Using a quite different dataset, Athavale and Edmister (2004) examine the pricing of a sequence of loans provided by the same bank in the U.S. This way they avoid using proxies for relationship strength. Based on OLS regressions, they find that the interest for follow up loans decrease with respect to the first loan. They interpret this result as support that lending relationships resolve information asymmetries between the bank and the borrower.

Some recent papers examine the effect of relationship lending on the simultaneous determination of various loan contract covenants. This approach allows incorporating the interdependencies between contract terms. Although a very appealing research direction, all of these studies are subject to an identification problem; which requires for instance, identifying instrumental variables that affect the determination of the interest rate but not collateral. Dennis, Nandy and Sharpe (2000) propose a four equation model for the interest rate, collateral, fees and maturity that is estimated for a sample of 2634 bank revolving contracts. Data comes from LPC Dealscan database. The proxy for relationship strength is loan concentration, defined as the amount of borrowings in the deal relative to the borrower's total debt. They find that interest rate increases as a relationship develops.

Brick and Palia (2005) use a simultaneous equation approach to account for the fact that collateral requirement is endogenously determined with interest rate. They find that the length of the relationship does impact upon both the probability of posting collateral and the level of the loan interest rates; however, the economic impact is relatively small.

D'Auria, Foglia and Marullo-Reedtz (1999) examine a panel dataset of Italian bank-firm relationships during the period 1987-1994, corresponding to 2300 large and medium-sized firms. They find that a main bank (measured as percentage of loans from main bank over total firm loans) provides credit at a lower interest rate and that increasing the number of bank relationships decreases the interest rate. Cosci and Meliciani (2002) also provide evidence from Italy. They find that the number of bank relationships has a positive effect on credit availability but has no effect on interest rates. With data of 18,000 loans supplied by one of the largest Belgian banks, Degryse and Van Cayseele (2000) find an increase in the interest rate and a decrease of collateral with the duration of relationship. All these studies are summarised in Table 2-3.

...........Table 2-3 goes around here...........

2.3 Default literature

In the main, the financial intermediation literature on small business lending mainly focused on the determinants and costs of credit access. There has been however little research examining the repayment behaviour of small firms that actually received loans. This was previously attributed to data limitations (Glennon and Nigro, 2005; 2008). To date, two strands of research in literature have focused on modeling default activity for small business ventures. The one strand has its foundation in financial ratio analysis based on financial statements and industry data (see for example studies by Altman and Sabato, 2007; Fidrmuc and Hainz, 2010; Behr, Guttler and Plattner, 2004 and Dyrberg-Rommer, 2005). The key motivation of these studies was to show the significant importance (for banks) of modeling credit risk for small firms separately from large firms. In general these studies build default predicting models based on financial ratios derived from large firms to determine whether these models can help predict default in small business ventures. In general, these studies find that models designed for large firms perform poorly in predicting small business default. They also show that a small number of financial ratios tailored to small firms namely: indebtedness, liquidity, profitability and sector-specific effects are important determinants of default. Hence, these studies conclude that banks should develop credit risk models specially addressed to small business ventures.

A key limitation has been identified in adopting the financial ratio analysis method when modeling small business default. Most of the small business ventures in these studies are actually ‘larger’ small firms (with sales of $50 million). In cases where financial statement data does not exist - typically early stage ventures such as business start-ups and young firms - financial ratio analysis technique cannot be applied; hence the problem of modeling default risk specific to small businesses remains.

A second strand of literature models small business default activity based on credit information of the principal business owner (see for example Berger *et al*, 2005; Agarwal *et al*, 2007; DeYoung *et al*, 2007; Berger *et al*, 2009). This strand of literature asserts that the personal credit history or indebtedness of small business owners is highly predictive of the loan repayments of their business. This is especially true for businesses at early stages of firm development, when business owners are usually the major (and probably the only) decision makers (Cassar, 2004).

Agarwal *et al* (2007) study the impact of borrower credit scores vs. business credit scores in predicting small business default risk. They find support for modeling small business risk based on credit information of the business owner; they show that business owners’ personal credit scores lower loan default vs. business credit scores.

Berger *et al* (2005) examine the effects of credit rating on the availability, price and default risk of small business credits. They compare small business loans granted before adoption of credit rating to those granted after the adoption of credit rating in managing credit risk. They find that the adoption of credit rating is associated with expanded quantities, higher average prices, and greater default risk for small business loans. One explanation that Berger *et al* (2005) put forward for their observation is that adopting credit rating as a mode of predicting default risk expands credit to some relatively risky 'marginal borrowers' that would otherwise not receive credit. The study by Berger and Frame (2007) compare banks that have adopted the use of credit rating to manage default risk to those who have not. They find that banks that use credit rating tend to have no more loan performance problems than other banks, despite the observed increase in lending to presumably more marginal borrowers.

It is plausible that lenders could be experienced (hence efficient) in appraising default risk, such that credit rating is not necessarily a key driver of default; lenders may opt to compensate for default risk by simultaneously adopting other instruments, such as requesting collateral that can be seized in the event of default (Berger and Udell, 1995) or by charging higher interest rates (Stiglitz and Weiss, 1981). Indeed theoretical papers by Bester (1985), Chan and Kanatas (1985), Besanko and Thakor (1987) and Chan and Thakor (1987) assert that collateral pledged helps align the interests of both lenders and borrowers, avoiding a situation in which the borrower makes less effort to ensure the success of the project for which finance was given. Hence these studies predict a negative relationship between the provision of collateral and loan default. It is also likely however that by the mere presence of collateral, lenders might apply less rigor when appraising loans that offer collateral; thus leading to an increase in default risk.

Other studies in the literature predict that the use of soft qualitative information, such as firm age and information gathered through contact over time (based on relationships), attenuates adverse selection and moral hazard issues; hence prove to be critical in loan approval decision-making and loan-pricing processes (Cole, 1998; Petersen and Rajan, 1994; Berger and Udell, 1995). It follows then to ask whether firm age and relationship variables help predict default risk.

Previous research has shown that firm age has a strong influence on loan default. For example, Bates and Nucci (1999), Evans (1987) and Dunne, Roberts and Samuelson (1989) have found that in the U.S, the probability that a firm will fail over a given period of time decreases with firm age. Studies by Good and Graves (1993) and Honjo (2000) state that new firms fail at higher rates than established firms (all else equal); which leads to the supposition that new firms are statistically more likely to default than established firms. Cowling and Mitchell (2003) also find that start-up businesses had higher default rates. They report that over time however start-ups were no more likely to default than existing firms. Glennon and Nigro (2005) also report that borrowers that are less than 3-year old at loan origination were more likely to default on their loans than more mature small businesses. Larger firms were shown to be more likely to default (Cowling and Mitchell, 2003; Glennon and Nigro, 2005, 2008).

It remains vague however whether relationships do indeed lead to loans that performs better in terms of default. Matteo *et al* (2012) report that the acquisition of new customers about whom less information is available - adverse selection mechanisms and the potential reduction of selection criteria - produce a higher level of bad debts for bank loans. This evidence emphasises the role of soft information and knowledge of the borrower in terms of strategies and intermediation models that consistently try to obtain and maintain the high quality of loan portfolios. Jimenez and Saurina (2004) tested factors driving default for 3 million bank loans during the period 1988–2000. Their analysis shows that as the number of banking relationships grows the probability of default decreases.

Finally, Studies by Cowling and Mitchell (2003); Glennon and Nigro (2005, 2008) observe significant effects on default risk across industrial sectors. Glennon and Nigro (2005; 2008) note that small firms in the manufacturing and retail sectors had significantly higher default rates and that firms in the transport and communications sector had significantly lower default rates than all other sector. Similarly, Cowling and Mitchell (2003), based on UK data, also find that loans to firms in the retail sector are more likely to default (all else equal). They further report that firms in the service sector are less likely to default relative to firms in other industries - a result that reflects the higher default rate for retail firms and the lower default rate for service firms.

To sum up the discussion so far, we have reviewed literature on the determinants of small business loan credit allocation, cost of credit and factors driving loan default. We now move on to discuss these key issues within the P2P lending context in the next session.

2.4 P2P lending

The concept of P2P lending is fairly recent and novel; hence literature on the topic in general, and specific to small business ventures, is scant. Since research in this area is relatively new, the concept of P2P lending itself does not enjoy a unique definition in the literature. Therefore we start in section 2.4.1 by defining the concept of P2P lending, and then we identify the benefits and costs of P2P lending. Next, in section 2.4.2 we briefly review some of the available literature concerned with P2P lending and small businesses finance.

2.4.1 Defining P2P lending

In this dissertation, P2P lending is defined as a way of attracting funds directly from multiple investors using the internet. Individuals lend money to others directly, without intermediation of traditional financial institutions like banks. Lending takes place predominantly through websites (known as platforms). It is typically a form of unsecured debt finance, rather than equity. P2P lending in this context is not restricted to specific relationships such as entrepreneurs lending (or borrowing) funds to other entrepreneurs. In addition, P2P lending is not restricted between individuals who share similarities in terms of wealth, gender, race or ethnicity transacting with each other. Instead, P2P lending explicitly defines a lending mechanism such that any individual may transact with any other individual over P2P lending websites.

As illustrated in Figure 2-1, P2P lending is a subset of crowdfunding. In one of the few published overviews of the topic, Mollick (2013: 3) defines crowdfunding within an entrepreneurial context, as “the effort by entrepreneurial individuals and groups (cultural, social; for-profit) to fund their ventures by drawing on relatively small contributions from relatively large number of individuals using the internet, without standard financial intermediaries”. There are four main contexts in which individuals fund projects on crowdfunding websites as illustrated in Figure 2-1 namely: donations based context, rewards based context, lending based context and equity based context (Mollick, 2013; Massolution, 2013). The fundamental distinction between these four types of providing finance is the lender’s intention and/or expectations concerning returns.

...........Figure 2-1 goes around here...........

The *donations based* crowdfunding follows a patronage model placing lenders in the position of philanthropists. Here lenders donate small loan amounts to small business ventures in economically underdeveloped regions in the world; and expect no direct returns for their efforts. One such example is Kiva.com (the largest donations based crowdfunding website) which has reportedly disseminated to date just under $400 million for entrepreneurs in developing countries (Massolution, 2013). Lenders who engage in *rewards based* crowdfunding receive a (non-financial) reward for backing entrepreneurial projects. Rewards include whatever product the small firm is selling (or intends to sell); for example electronic gadgets, CDs, t-shirts, beer etc. (Massolution, 2013). Rewards based crowdfunding may also treat lenders as customers - allowing them access to the products produced by the funded project at an earlier date or at a cheaper price (Massolution, 2013). Mollick (2013) in one of the few studies published in crowdfunding asserts that pre-selling of products to earlier customers is a common feature of those crowdfunding projects that more resemble entrepreneurial ventures, such as producing novel software, hardware, or customer products. Kick-starter is the largest rewards based crowdfunding website; with over $380 million lent trough this site.

Individuals who engage in the *equity based* context and the *lending based* context do so with an intention of receiving a reasonable return as compensation for taking a risk of extending funds. As of mid-2013, however, equity based lending is generally not permitted in the United States and still relatively rare worldwide, making up less than 5 percent of all crowdfunding investment (Massolution, 2013). This is due to the fact that equity crowdfunding is subject to high levels of regulation (Heminway and Hoffman, 2010); hence the eventual adoption of this approach (relative to other forms of crowdfunding) is still uncertain. Finally, the *lending based* model is one in which funds are offered to potential borrowers as a loan, with the expectation of a rate of return on capital invested. Lendingclub.com and Prosper.com are the largest P2P lending websites in the world; they have originated over $2 billion unsecured loans. For the most part, P2P lending functions on the basis of trust, albeit trust between people that have only met on the internet. The sites match individual borrowers with individual lenders and provide lenders with both financial and nonfinancial information to facilitate due diligence. Borrowers post their loan requests on the websites, share (voluntary) information about themselves - both personal and financial - and lenders decide whether or not to contribute to their loan request. Every loan is underwritten by multiple individual lenders, each committing a fraction of the loan until it is funded in full. Once fully funded, the loan is originated and the lenders receive a pro rata share of the principal and interest payments until the loan reaches maturity or until the borrower defaults. The websites generate their revenue via service fees, which they collect from borrowers as well as lenders (Klafft, 2008). Borrowers are afforded an opportunity to include text description in their loan request, which lenders can utilise to make a compelling case for why they should extend credit to these borrowers. In addition, most P2P lending websites allow borrowers to include a picture when requesting for a loan.

Some of the benefits of P2P lending (over traditional lending institutions) for small business borrowers this includes include:

- Access to millions of individual lenders at one time - which increases the chance of receiving funding for small business ventures;

- Relaxation of collateral requirements since the loans are unsecured - which previous research has shown to be one of the reasons for persisting credit rationing (Evans and Jovanovic, 1989);

- Quick turnaround times of obtaining the loan - on average it currently takes around 7 days to get a loan on P2P lending websites with minimal paper work. For lenders - P2P lending provides better returns, which seem to outperform banking returns (albeit given relatively higher risk).

Moreover, given that lenders can diversify their investment to be as little as $50 per transaction; this limits the impact of loss. As with just about every financial model, there are also disadvantages to P2P lending as well; particularly for lenders. First, there is very little reassurance that investments will be paid back on time (if at all). Second, unlike traditional institutions, given that loans in this context are unsecured and can be as little as $50, there are few methods of recourse for non-payment. If a lender loses $50 it may not be worthy to chase the borrower.

2.4.2 P2P lending and small business finance

Despite the potential of P2P lending being a viable method of funding small business ventures, there has been very little published peer reviewed work to date on P2P lending and small business finance – apart from a small stream of studies focus solely on rewards based crowdfunding (Agrawal *et al*, 2011; Mollick, 2013; 2010; Schwienbacher and Larralde, 2012). In general rewards based lending studies are concerned with the assessment of quality of early stage entrepreneurial ventures based on signals of quality that venture capitalists typically look for in their selection process. For example, Agrawal *et al*, (2011), based on data from Kickstarter.com, used a market of musicians seeking crowdfunding to understand whether crowdfunding relaxes geographic constraints on fundraising that are typical of venture capital firms. Mollick (2010) uses data from Kickstarter.com to examine crowd funded projects that match characteristics of more traditional venture capital backed seed ventures to determine what role geography and gender play in new venture finance within a crowdfunding regime. Other studies examine aspects of efficient communication and networking as determinants of early stage venture funding success (Schwienbacher and Larralde, 2012). Finally, the study by Mollick (2013), also based on data from Kickstater.com, offers one of the first generic analyses of how rewards based crowdfunding works; sharing insight on the ways in which the characteristics of (potential) small business owners and the way they present their ventures can affect entrepreneurial financing outcome.

Looking specifically at P2P lending research, so far almost all existing literature is typically concerned with understanding P2P lending dynamics for general loans; investigating determinants of funding outcome on the basis of: perceived trustworthiness (Duarte *et al*, 2010; Klafft, 2008), taste-based discrimination (Pope and Sydnor, 2011; Ravina, 2008), borrowers' identity claims in the stories that they tell (Herzenstein *et al*, 2011; Sonenshein *et al*, 2011), and interest rates (Iyer *et al.*, 2010). Another stream of research examines the social network aspect of P2P lending, such as how social networks affect loan performance (Freedman and Jin, 2008), how social networks relate to loan default risk (Everett, 2010), and how the strength and verifiability of relational networks influence funding outcomes and loan defaults (Lin, Viswanathan and Prabhala, 2011).

In Tables 2.4 to Table 2.6 we review each of these studies in detail; highlighting what has been studies in terms of variables and what might still need further investigations. As shown in the review tables, In general these studies are based on large samples in the thousands, mostly gathered from Prosper based in the US; and include data related to a range of personal and background characteristics affecting the key dependent variables of interest. Moreover, a variety of operational definitions and analysis procedures are utilised in the different studies mostly developed from (linear) multivariate regression models. We found that a subset of studies concerned with understanding P2P lending dynamics for general loans included a binary variable in the analysis, where they compare funding outcome, or cost of finance, or default behavior between general loans and business loans (see for example see Ravini, 2008; Pope *et al*, 2011; Weib *et al.*, 2010; Barasinska, 2010 and Duarte *et al.*, 2010).

...........Table 2-4 to Table 2-6 goes around here...........

Ravini, (2010) was among the first to analyse P2P lending data. The working paper strived to determine the role of beauty and physical appearance on both credit extension and the cost of credit. They use data from Prosper; covering general loans over a one month period from March - April 2007. To analyse the data, liner Probit regressions were used to estimate credit allocation and OLS regressions were used to estimate factors driving interest rate. The study includes a variable to compare the funding success of entrepreneurs in this context based on the self-reported employment status of each borrower. They report that borrowers who are full time entrepreneurs are more likely to be funded when compared to their counterparts who report that they are in full time employment. They find a positive but insignificant result for the interest rate estimation.

Pope and Sydnor, (2011) in a published paper, analyse Pictures to determine the role of discrimination in affecting access to credit, cost of credit and loan default behaviour. The study is based on sub sample of data from Prosper; covering general loans over a one year period from 2006 to 2007. To analyse the data they use liner Probit regressions to estimate credit allocation; OLS regressions to estimate factors driving interest rate and a Cox hazard model to estimate default. In their analysis, Pope and Sydnor (2011) include a control variable to compare the risk of general loans in relation to that of business loans. They report that small business loans are less likely to be funded, are more likely to pay higher interest rates and are more likely to result in default. These results cannot be generalised however to the whole population of loans on Prosper, as they are based only on a subsample of loan requests that include pictures. Inclusion of pictures is optional; hence there results may be biased due to selection effects.

The working paper by Weib *et al* (2010) analyses the impact of non-verified information vs. verifiable information in determining both credit extension and the cost of credit. They hypothesise that the verification of certain borrower characteristics significantly affects funding success and interest rates paid for the loans. The study is based on data from Prosper; covering general loans over a three month period; from July to October 2009. To analyse the data, they use a multinomial logic regression (credit allocation) and a standard OLS regression for the interest rates estimation. In both estimations, the business variable is insignificant.

Duarte *et al*, (2010) analyse the role of trust in financial decisions; concerned with assessing the extent to which people in general use impressions of potential counterparties' trustworthiness when making financial decisions. They construct the measure of trustworthiness based on the pictures included in the loan requests. The study is also based on data from Prosper; covering general loans over from 2006 to 2009. They use liner Probit regression to estimate credit allocation and Cox hazard model to estimate default. They find that business owners s who are perceived to be less trust worthy are less likely to have their loan request funded; however they find that their loans are no more likely to default than those who are perceived to be less trustworthy. More specifically, they find that borrowers who include business premises in their loan request seem to be perceived to be more trust worthy - hence they are likely to be funded.

Barasinska (2010) models the role of gender in affecting credit access. The study is based on data from Smava, a German P2P lending company; covering general loans over a one year period from 2006 to 2007. They use liner probability models to estimate credit allocation. Similarly, this study reports that small business loans relative to general loans are more likely to be funded.

Overall, given the diverse operational definitions and procedures adopted, as to be expected, we found a number of conflicting results. That is, there were a few cases where one study finds significant positive effect between general loans and business loans and the other a significant negative effect. Some of the observed patterns are non-complementary. These findings are summarised in Tables 2-4 to Table 2-6.

2.5. Research gap

We summarise our discussion of the literature by illustration with reference to Table 2-7. Starting at the top left corner of Table 2-7 and moving counter-clockwise, the literature on traditional lending offers a broad understanding on some of the big questions asked by researchers in entrepreneurial finance, which help set the agenda of this dissertation. This literature also offers theory on mechanisms typically used to lessen information asymmetries; which we seek to test in this new context. It is important to understand whether information asymmetries issues are attenuated by the same underlying mechanisms as other forms of traditional lending or if there are some other mechanisms adopted. This is an important topic because it has yet to be established at an empirical level whether P2P lending merely crowds out traditional small business finance or whether it actually provides finance to businesses that otherwise would not get it.

...........Table 2-7 goes around here...........

Moving to the bottom left hand corner of Table 2-7, the literature on general loans from traditional institutions does not inform this study.

Next, the broad literature on P2P lending is mostly concerned about general loans. A subset of this literature however includes a binary variable to distinguish credit risk of business loans in comparison to general loans. Although this literature offers valuable contributions in understanding a new way of mobilising finance; none of the reviewed studies code for business loans specific data. Hence, little is known about the type of business ventures that come to this market for funds as well as the dynamics governing successfully raising funding in this context. We also don’t know whether this new form of credit favours specific industries; or even whether this market is sustainable in the long run. Consequently, we have collected and created business loan specific data; introducing new key variables to help us study small business loans within the

P2P lending context. Hence this positions our research in the top right corner of Table 2-7. Since P2P lending is novel and potentially disruptive to traditional approaches of funding small business ventures, we follow entrepreneurial finance literature and P2P lending literature by formulating the following key research questions:

- What factors drive the probability of funding for small business loans in the P2P lending context?
- What factors drive the cost of credit for small business loans in the P2P lending context?
- Looking at the default activity in comparison to lender returns, is it worth it for lenders to extend small business loans in the P2P lending context?

Collectively, the first two questions will help determine whether P2P lending merely crowd out traditional banking finance or whether it offers finance to those who are traditionally underserved; whilst the third question will help determine whether P2P lending is sustainable in the long run.

In order to operationalise these research questions, the basic intuition behind the analysis is based on the fact that due to the novelty of P2P lending, currently there are no established mechanisms of coping with asymmetric information that have been identified among lenders financing small business ventures in this context. We test the proposition that if P2P lending responds to known signals established in traditional lending, it reinforces both the validity of these mechanisms in predicting borrower quality and the ability of 'amateur' lenders to select small businesses risk; which may render the market somewhat viable in extending credit to small business ventures. If on the other hand the mechanisms used by traditional lending institutions are not predictive of borrower quality and the ability of lenders to select small businesses risk in P2P lending, this may suggests that either the signals sought by traditional lenders are ceremonial in nature and are not generally predictive of entrepreneurial credit allocation outside the traditional 'offline' lending market; or that P2P lending relies on other (unique) mechanisms.

Additionally, in traditional lending, many of the mechanisms available to already established firms may not necessarily be available to new business start-ups. Hence, we test the proposition that if mechanisms used to cope with information asymmetries are no different for new business start-ups when compared to established businesses in the P2P lending context, perhaps P2P lending may be introducing small business finance to a new market of business ventures; previously under-served. This is a particularly important distinction as new firms are widely perceived to be the riskiest class of small business ventures (Bates and Nucci, 1999; Evans, 1987; Dunne, Roberts and Samuelson 1989).

In terms of the data aspects, when reviewing P2P lending studies we found that almost all empirical studies on loan performance were based on loans that were current, given that the loans had not yet reached full maturity during the period of study (with the exception of Duarte *et al*, 2010). In many settings, there is no *ex post* performance data. Consequently, these studies operationalised default based on the assumption that all loans which were delinquent or late (and not current) would eventually result in default (Pope and Sydnor, 2011; Ravini, 2008; Everett 2010; Kumar, 2007). Unlike the reviewed studies, our data set incorporates loan performance data that includes at least three cycles of loans which have reached full maturity. This enables us to model loan performance founded on actual default observed. Hence, given the availability of this loan performance data, we were able to validate whether 'amateur' lenders were indeed efficient (or not) in appraising credit risk. We test the proposition that investing in small business loans associated with higher (lower) risk will manifest in poorer (greater) returns for lenders. We also test whether the proposition puts forward by theory that new firms are more risky than established firms is supported by loan performance data in the P2P lending context..

In terms of methodology, we found that in almost all reviewed P2P lending studies, the estimation of the interest rate (OLS) and default equations (Probit) are based on funded loans selected non-randomly from the sample of all loan requests. If selectivity exists, the coefficients observed from the estimations may be biased (Greene, 1997). Therefore, we follow these studies however, we also account for selectivity by adopting 2 stage models (Cressy, 1996). Finally, all reviewed P2P lending studies employ linear multivariate models. We follow these studies and adopt a similar approach in our empirical analysis. However, unlike the P2P lending studies, we also account for non-linearity of some of the key variables of interest (Cole, 1998; Cowling, 1997; Burke and Hanley, 2006).

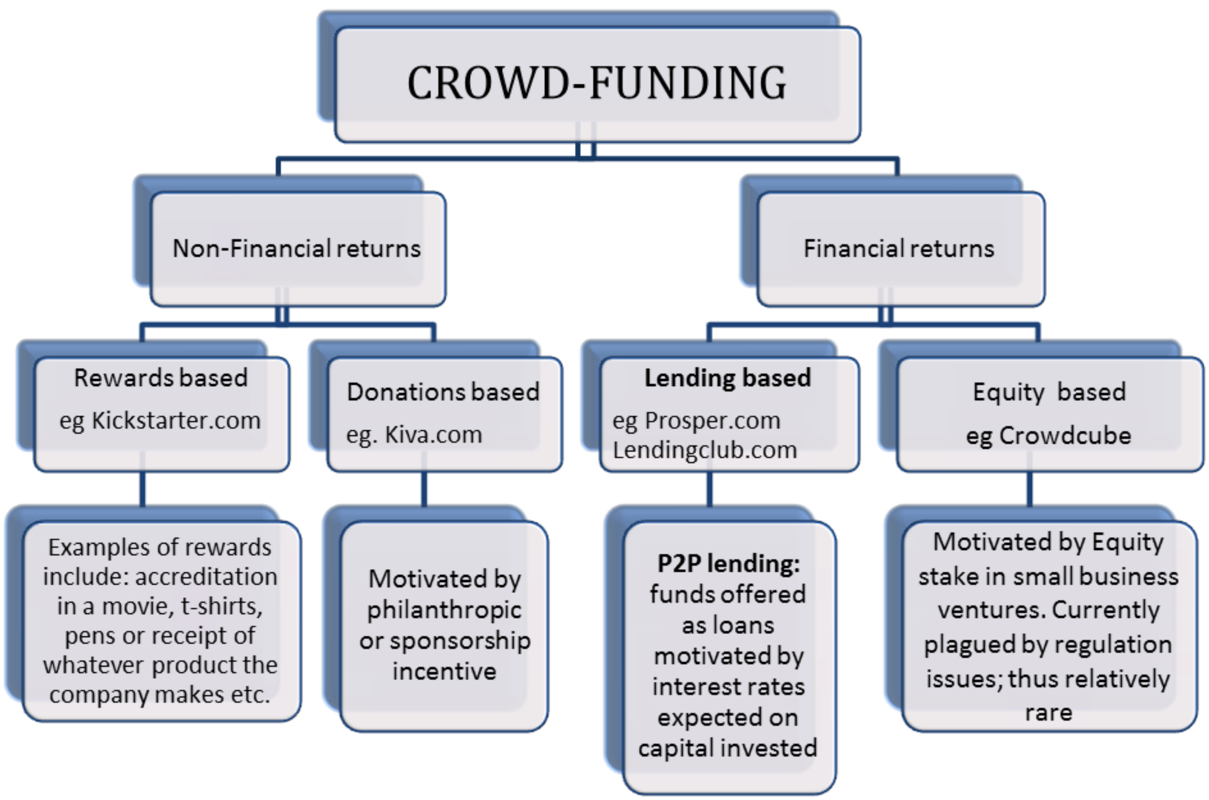
Lastly, we follow P2P lending literature and control for non-verified information in the form of text elaborations (Weib, 2010; Herzenstein *et al*, 2011; Sonenshein *et al*, 2011); borrower's labour force activity (Ravini) and pictures (Pope and Sydnor, 2011; Ravini, 2008; Duarte *et al.*,2010). We also control for geographic location (Agrawal *et al*, 2011).

2.6 Chapter summary

The entrepreneurial finance literature offers a broad understanding on some of the big questions asked by researchers in entrepreneurial finance, which help set the agenda of this dissertation. This literature also offers theory on mechanisms typically used to lessen information asymmetries. The broad literature on P2P lending is mostly concerned about general loans. Although this literature offers valuable contributions in understanding a new way of mobilising finance; none of the reviewed studies code for business loans specific data. The lack of empirical studies specific to small business ventures and the availability of the data from P2P lending websites motivated us to investigate the impact of P2P lending on the finance of small business ventures.

Figure 2-1: P2P lending as a subset of Crowd-funding typology

P2P lending is one of four common forms of attracting small business capital from a group of strangers encompassed under the crowd – funding family. Lenders extending capital to small business ventures through P2P lending are motivated by interest rates expected on capital invested.



Source: Adapted from Massolution, 2013

Table 2-1: The overarching research issues in small business finance

This table highlights the key issues and related smaller questions motivating the small business finance research agenda

The big issues	Related smaller questions
(1) **Credit allocation:** *Which firms get funded which firms don't*	• Why do banks routinely deny credit rather than charge higher borrower prices • How does collateral affect credit allocation • Is there ethnic discrimination in the market for small business credit • What is the role of information and bank- relationships on credit allocation • Is there gender discrimination in the market for small business credit • Can credit contract innovations resolve the problems that bring about credit rationing • Is credit rationing a static or dynamic phenomenon
(2) **The pricing/cost of small firm finance:** *What are the determinants of loan premium (or interest rates) for small business credit*	• What are the effects of collateralisation on bank loan premium • What is the role of information and bank- relationships on bank loan premium • Is second round loan finance cheaper or costly for small business venture
(3) **Determinants of small business default**	• What are the determinants of loan default; • What is the importance of owner and business credit risk characteristics in determining default behaviour of small business loans • What factors determine default probability over time • Is there a difference in default rates in different sectors

Source: Compiled by author through literature research

Table 2-2: Theoretical and empirical studies of asymmetric information and Credit allocation in small business finance

This table summarises the main findings in the theoretical and empirical literature on the risk of asymmetric information and credit allocation to small business ventures. For the empirical studies we further describe the data source, sample size and the adopted method of analysis.

Theoretical Studies

Author	Summary of findings
Bester (1985)	Bester concludes that collateral can help sort a priori indistinguishable borrower. Bester concludes that the low risk loan applicants try to differentiate themselves from high risk applicants by accepting higher collateral as collateral is costly.
Bester (1987)	Studies collateral and its impact on moral hazard. Bester concludes that an increase in interest rates results in a negative effect over the repayment probability; whereas an increase in collateral results in a positive effect, i.e. an increase in collateral makes a risky project less attractive
Chan and Kanatas (1988)	They also conclude that collateral can help sort *a priori* indistinguishable borrowers
Besanko and Thakor (1987)	Examines the role of collateral in diminishing credit rationing when lenders do not know lender's default probability. In this paper the authors study loan contracting under asymmetric information within a multidimensional pricing scheme (i.e. they look at collateral, interest rates and loan quantity) and the possibility of rationing. In both cases, Besanko and Thakor find a positive relationship between borrower quality i.e. low quality borrowers (high risk) put up less collateral than high risk (low quality borrowers).

Empirical Studies

Author	Sample Description	Data Source	Method of analysis	Summary of findings
Cressy (1996)	Sample of 2048; period: 1988	UK start-up businesses	Two stage regressions	Typical firm unconstrained by finance, instead by human capital. They find a that older borrowers, with more work experience, who are educated are more likely to have access to collateral and therefore are more likely to have access to credit
Harhoff and Körting (1998)	1509 German SMEs	Survey data	Regressions	This paper finds that collateral decreases with relationship duration. They also find that the availability of credit decreases with number of lenders
Petersen and Rajan (1994)	3404 sample of firms with fewer than 500 employees	1987 US National Survey of Small Business Finance	Regressions	There is a positive relationship between the length of pre-existing relationships and access to credit. If the borrower has multiplier suppliers of finance the availability of finance will decrease

Table 2-2 continues

Author	Sample Description	Data Source	Method of analysis	Summary of findings
Empirical Studies				
Berger and Udell (1990)	1 million commercial loans from 460 US banks: period:1977-1988	Federal Reserve's Survey of Terms of Bank lending	Regression; cross section	They test Bester (1985)'s theory that better risk borrowers pledge collateral by comparing secured vs. unsecured loans. This paper finds that contrary to Bester's theoretical model, Collateral associated with higher credit risk. Their results suggest that borrowers who pledge collateral are riskier than borrowers who do not. As evidence of this – they show that borrowers with secured loans vs. unsecured loans tend to have more non-performing loans (delinquencies, re-negotiated etc. Berger and Udell, 1990: 40). This correlation can be explained in the context of symmetric information (in the case of the bank as mentioned by de Meza), with the entrepreneur's over optimism about their projects
Boot , Berger and Udell (1991)	Proof of Chan and Kanatas model that collateral resolves private information and moral hazard problem	Federal Reserve's Survey of Terms of Bank lending	Regressions	Their analysis generalises the results of the papers by Bester (1985); Chan and Kanatas (1987); to the case of both moral hazard and adverse selection. This study thus explains collateral variations within observationally indistinguishable group of borrowers as well as cross-sectionally distinct groups. They find that higher quality borrowers do not necessarily post more collateral.
Hanley and Girma (2006)	466 new ventures. period: 1998 - 1999	UK retail bank (identity not disclosed)	2 stage regression	Report a positive and significant relationship between collateral and access of credit. They also report that lenders reject borrowers looking for smaller loans – which they interpret as the possibility that lenders are likely to reject lower credit requests if these projects are synonymous with underinvestment and hence pose a higher risk of failure. This low success rate is reflected in the higher interest margins these loans are associated with.
Berger and Udell (1995)	3400 sample of firms with fewer than 500 employees	1987 US National Survey of Small Business Finance	Regressions	Test Boot and Thakor (1984) theory on bank lender relationship. Evidence in this paper shows that small firms with longer banking relationships are less likely to pledge collateral relative to other small firms. These results suggest that banks accumulate information over the length of the relationship that is used to reduce information asymmetries. Therefore for small firms, the alleviation of collateral requirements increases access to credit.
Cole (1998)	Nationally representative sample of 5356 small businesses; period: 1991-1994	1993 US National Survey of Small Business Finance	Logit Regressions	This paper extends Petersen and Rajan (1994), and puts an emphasis that it's the existence of a relationship (and not necessarily the length of relationship) that influences credit allocation. The paper also finds that multiple borrowing has a negative on the availability of credit, along with firm age, business/personal delinquencies and business/personal judgements

Table 2-3: Theories and evidence of the factors driving interest rates for small business loans

This table summarises the studies that look at factors that drive the cost of finance for small business ventures. The table highlights studies that put forward the theories underpinning the predicted relationships between interest rates and collateral, interest rates and pre-existing relationships as well as theories underpinning the predicted relationship between interest rates and access to follow up loans.

	Theories underpinning the studies	Empirical evidence which support this correlation
Predicted relationship between ***interest*** rates and ***collateral***		
Negative	Bester(1985); Besanko and Thakor (1987)	Cressy (1996); Cressy and Toivanen (2001); Keasey and Watson (2000) Cowling,(1999)
Positive	De Meza and Southey (1996)	Berger and Udell (1990)
Convex U – Shape	Cressy(2000), Burke and Hanley (2003)	Burke and Hanley (2006)
Predicted relationship between ***interest*** rates and ***pre-existing relationships***		
Negative		Berger and Udell (1992); Berger and Udell (1995)
No relationship		Petersen and Rajan (1994); Harhoff and Körting (1998)
Predicted relationship between ***interest rates*** and ***second period loans***		
Positive	Greenbaum *et al*, (1989); Sharpe (1990); Wilson (1993)	Hanley and Crook (2005); Binks and Ennew (1998)
Negative	Diamond (1989); Boot and Thakor (1994); Petersen , Rajan and Raghuram (1995)	

Table 2-4: Analysis of factors affecting credit allocation on P2P lending websites

This table summarises the main findings in the P2P lending empirical literature on determinants driving credit allocation. We report the sign of the effect of the respective factors on credit allocation as positive and significant +; positive and insignificant (+); negative and significant -; variable not reported ? And variable not included x.

Probability of funding estimation: $Y = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Herzenstein *et al* (2008)	Kumar (2007)	Lin *et al* (2011)	Freedman *et al* (2008)	Herzenstein *et al* (2011)	Herzenstein *et al* (2011)	Barasinska (2011)	Sonenshein *et al* (2011)	Duarte *et al* (2008)	Weib *et al* (2010)	Ashta *et al* (2009)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Not reported ? New variable: (red) Empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) All loans Jun 06 – May 07	Prosper.com (US) general loans Mar 07 – Apr 07	Prosper.com (US) general loans June 2006	Prosper.com (US) general loans July 07 -- Dec 07	Prosper.com (US) general loans Jan - May 08	Prosper.com (US) general loans Jun 06- Jul 2010	Prosper.com (US) general loans Jun06 – Jun 07	Prosper.com (US) A general loans June 2006	Smava.com (Germany) All data Mar 07-Mar 10	Prosper.com (US) general loans June 06	Prosper.com (US) general loans May 06 – Jan 08	Prosper.com (US) general loans Jan 07 – May 08	Prosper.com (US) general loans Jan 07 – May 08
Variables													
Listing Characteristics													
# listings before current	x	-***	x	x	x	x	x	x	x	x	x	x	x
$ amount requested	- ***	-***	-***	-*	-***	-*	-***	-*	-***	-***	-**	-*	- ***
($ amount requested)[2]													
0ffer Interest rate	+ ***	+***	+***	+*	+***	+***	+**	+**	+***	+***	+**	+**	+ ***
(offer interest rate)[2]													
# words - loan description	x	x	+*	(+)	+***	+*	+***	x	+***	+*	x	+*	x
Close when funded	x	-***	x	x		x		x	x	x	-**	x	x
Auction Format	x	x	x	x	+***	x	+***	x	x	x	x	x	+***
Loan Characteristics													
Loan Duration (60mnths)	x	x	x	x	x	x	x	x	-***	x	x	x	x
Business loans (relative to general loans)	- ***	+**	-**	x	x	-*	(-)	x	+**	-**	+*	(-)	x
Loan Purpose													
Working Capital													
Start-up costs													
Purchasing building													

Table 2-4 continues

Table 2-4 continues

# lenders funding loan	x	x	x	x	x	x	+***	x	x	x	x	x	+***
Borrower Characteristics													
Debt to Income	- ***	x	-***	(-)	-***	-**	-***	-*	x	-***	-***	_***	-*
Bankruptcy (last 10 yrs.)	?	-**	x	x	x	x	x	x	x	x	x	_***	x
Bankruptcy (last 2 years)	?	x	-*	-*	x	x	-*	x	x	-*	x	_***	x
Bankruptcy (last 12m)	?	(+)	X	x	x	x	x	x	x	x	x	_***	x
Ever default (last 7 years)	?	-**	-*	-**	-**	x	x	-*	x	-*	-**	_***	x
# credit lines	?	(-)	(-)	x	x	x	x	(-)	x	(-)	X	_*	x
Revolving credit Balance	?	(-)	(-)	x	x	x	x	(-)	x	(-)	-**	_*	x
Revolving Credit amount	?	x	(-)	x	x	x	x	(-)	x	(-)	-**	_*	x
Verified bank account	?	+***		+**	+***	x	x	x	x	x	x	x	x
Credit Rating													
AA	+ ***	+ ***	+ ***	Base	Base	+ ***	+ ***	+ ***	Base	+ ***	+ ***	+ ***	+ ***
A	+ ***	+ ***	+ ***	-**	-***	+ ***	+ ***	+ ***	-***	+ ***	+ ***	+ ***	+ ***
B	+ ***	+ ***	+ ***	-**	-***	+ ***	+ ***	+ ***	-***	+ ***	+ ***	+ ***	+ ***
C	+ ***	+ ***	+ ***	-**	-***	+ ***	+ ***	+ ***	-***	+ ***	+ ***	+ ***	+ ***
D	+ ***	+ ***	+ ***	-**	-***	+ ***	+ ***	+ ***	-***	+ ***	+ ***	+ ***	+ ***
E	+ ***	+ ***	+ ***	-**	-***	+ ***	+ ***	+ ***	-***	+ ***	+ ***	+ ***	+ ***
HR	Base	Base	Base	-**	-***	Base	Base	Base	-***	Base	Base	Base	Base
Borrower Income	?	x	+*	x	x	x	x	+ ***	x	+*	+*	+*	+*
Borrower Income range													
$25,000- $49,999	?	+ ***	x	x	x	x	+ ***	+ ***	x	x	x	x	x
$50,000- $74,999	?	+ ***	x	x	x	x	+ ***	+ ***	x	x	x	x	x
$75,000- $99,999	?	+ ***	x	x	x	x	+ ***	+ ***	x	x	x	x	x
$100,000+	?	+ ***	x	x	x	x	+ ***	+ ***	x	x	x	x	x
Employment Status													
Self employed	?	(+)	x	x	x	x	(+)	(+)	(-)	x	x	(-)	x
Employed part time	?	(-)	x	x	x	x	(-)	(-)	X	x	x	(-)	x
Unemployed	?	+**	x	x	x	x	+**	+**	-***	x	x	+*	x
Retired	?	+***	x	x	x	x	+***	+***	-***	x	x	(-)	x
Length of Employment Status	?	(+)	x	x	x	x	(+)	(+)	x	x	x	(+)	x
Home Ownership	?	+**	+	+	x	x	+**	+**	x	+	x	(+)	x
Education	X	0	x	x	x	x	0	0	x	x	x	x	x
Picture													
Gender Male (base group)	base	base	base	x	x	x	base	x	(-)	base	x	(-)	x
Female	+ ***	(+)	(+)	x	x	x	(+)	x	base	(+)	x	base	x
Age (base group 35yrs – 60 yrs)	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 2-4 continues

Young(less than 35yrs)	+ *	(-)	(+)	x	x	x	(-)	x	x	(+)	x	x	x
Old (more than 60)	- *	(-)	(-)	x	x	x	(-)	x	x	(-)	x	x	x
Race White (base group)													
Black	- ***	(-)	(-)	x	x	x	(-)	x	x	(-)	x	x	x
Hispanic	(-)	(-)	(-)	x	x	x	(-)	x	x	(-)	x	x	x
Asian	(+)	(+)	(+)	x	x	x	(+)	x	x	(+)	x	x	x
Beauty													
Attractiveness/Beauty	x	x	x	x	x	x	x	x	x	x	x	x	x
Very attractive	(+)	+**		x	x	x	+**	x	x	x	x	x	x
Very unattractive	(-)	(-)	x	x	x	x	x	x	x	x	x	x	x
Happiness/smile	x	x	x	x	x	x	x	x	x	x	x	x	x
Happy	(+)	(-)	x	x	x	x	(-)	x	x	x	x	x	(+)
Unhappy	- *	base	x	x	x	x	x	x	x	x	x	x	- *
Weight	x	x	x	x	x	x	x	x	x	x	x	x	x
Somewhat overweight	(+)	+***	x	x	x	x	+***	+***	x	x	x	x	x
Very overweight	-*	(-)	x	x	x	x		x	x	x	x	x	x
Professionally dressed	(+)	(+)	x	x	x	x	(+)	x	x	x	x	(+)	x
Signs of Military involvement	+ ***	x	x	x	x	x	x	x	x	x	x	x	x
Group Participation													
Belong to a group	x	x	+**	-	+***	+**	x	x	x	+**	x	+*	x
Religious	x	x	x	x	+***	x	x	x	x	x	x	x	x
Alumni	x	x	x	x	+***	x	x	x	x	x	x	x	x
Geography	x	x	x	x	+***	x	x	x	x	x	x	x	x
Group leader recommendation	x	x	+	+	+**	+**	x	x	x	+	x	+*	x
Group Leader funding	x	x	+***	+	+**	+**	x	x	x	+***	x	x	x
Group Size	x	x	x	x	-**	x	x	x	x	x	x	x	x
Firm Characteristics													
Industry SEC codes													
Age of Business (new firm or existing firm)													

Table 2-5: Analysis of factors affecting cost of credit on P2P lending websites

This table summarises the main findings in the P2P lending empirical literature on determinants driving the cost of credit. We report the sign of the effect of the respective factors on credit allocation as positive and significant +; positive and insignificant (+); negative and significant -; variable not reported ? And variable not included x.

P2P Related Studies: Effects of borrower and loan characteristics on the final interest for funded loans - OLS $I_{Final} = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Berger *et al* (2009)	Everett (2010)	Kumar (2007)	Lin *et al* (2011)	Iyer *et al* (2008)	Barasinska (2010)	Herzenstein *et al* (2008)	Herzenstein *et al* (2011)	Barasinska (2011)	Weib *et al* (2010)	Freedman *et al* (2008)	Duarte *et al* (2008)	Chemin *et al* (2010)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) General loans Jun 06 – May 07	Prosper.com (US) General loans Mar 07 Apr 07	Prosper.com (US) General loans Nov 05 – Aug 07	Prosper.com (US) General May 06 – 7 May 08	Prosper.com (US) General loans July 07 -- Dec 07	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Feb 2007 – Oct 2008	Smava.com (Germany) General data Mar07-Mar 10	Prosper.com (US) General loans June 2006	Prosper.com (US) General loans Jun06 – Jun 07	Smaca.com (Germany) General loans Mar07-Mar 10	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Jun 06- Jul 2010	Prosper.com (US) General loans May 06 – Jan 08	MyC4 (Africa – Uganda ,Kenya ,Ghana ,Ivory Coast)
Variables															
Listing Characteristics															
# listings before current	?	(-)	x	x	x	x	x	x	x	x	x	x	x	x	x
$ amount requested	+***	+***	+***	+***	+**	+**	+**	+**	+**	+**	+***	+**	+**	+***	+**
($ amount requested)^2															
Offer interest rate															
Default loan (ex post - binary 0 or 1)															
Close when funded	?	+***	x	+***	x	+**	x	x	x	x	x	x	x	x	x
Auction	?	x	-***	x	x	x	x	x	-**	x	x	x	x	x	x
Loan Characteristics	x	x	x	x	x	x	x	x	x	x	x	x	x	x	
Business loans (relative to general loans)	+*	x	x	x	x	x	x	x	x	x	x	(+)	(+)	x	x

Table 2-5 continues

P2P Related Studies: Effects of borrower and loan characteristics on the final interest for funded loans - OLS $I_{Final} = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Berger *et al* (2009)	Everett (2010)	Kumar (2007)	Lin *et al* (2011)	Iyer *et al* (2008)	Barasinska (2010)	Herzenstein *et al* (2008)	Herzenstein *et al* (2011)	Barasinska (2011)	Weib *et al* (2010)	Freedman *et al* (2008)	Duarte *et al* (2008)	Chemin *et al* (2010)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) General loans Jun 06 – May 07	Prosper.com (US) General loans Mar 07 Apr 07	Prosper.com (US) General loans Nov 05 – Aug 07	Prosper.com (US) General May 06 – 7 May 08	Prosper.com (US) General loans July 07 -- Dec 07	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Feb 2007 – Oct 2008	Smava.com (Germany) General data Mar07-Mar 10	Prosper.com (US) General loans June 2006	Prosper.com (US) General loans Jun06 – Jun 07	Smaca.com (Germany) General loans Mar07-Mar 10	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Jun 06- Jul 2010	Prosper.com (US) General loans May 06 – Jan 08	MyC4 (Africa – Uganda ,Kenya ,Ghana ,Ivory Coast)
Loan purpose															
Working Capital															
Start-up costs															
Purchasing Building															
Borrower Characteristics															
Debt to Income	?	x	+***		+**	+*	+**	+**	+**	+**	+**	+**	+**	+**	+*
Bankruptcy (last 2months)	?	+**	x	x	x	x	(+)	(+)	(+)	(+)	(+)	(+)	(+)	x	x
Bankruptcy (last 10 yrs.)	?	(+)	x	x	x	x	+***	+***	x	x	+**	+***	x	x	x
Ever default (last 7 years)	?	+***	x	x	x	x	+***	+***	+***	x	x	+***	x	x	x
# credit lines	?	+***	x	x	x	x	(+)	(+)	(+)	(+)	(+)	(+)	x	+***	x
Revolving credit Balance	?	(+)	x	x	x	x	x	x	x	x	x	x	x	x	x
Revolving Credit amount	?		x	x	x	x	(-)	(-)	(-)	(-)	(-)	(-)	x	x	x
Bank card utilization rate	?	(-)	x	x	x	x	+**	+**	+**	+**	+**	+**	x	x	x
Credit Rating															
AA	?	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	Base	- ***	- ***	base	x
A	?	-***	-***	-***	-***	-***	-***	- ***	- ***	- ***	+***	- ***	-***	+***	x
B	?	-***	-***	-***	-***	-***	-***	- ***	- ***	- ***	+***	- ***	-***	+***	x

Table 2-5 continues

P2P Related Studies: Effects of borrower and loan characteristics on the final interest for funded loans - OLS $I_{Final} = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Berger *et al* (2009)	Everett (2010)	Kumar (2007)	Lin *et al* (2011)	Iyer *et al* (2008)	Barasinska (2010)	Herzenstein *et al* (2008)	Herzenstein *et al* (2011)	Barasinska (2011)	Weib *et al* (2010)	Freedman *et al* (2008)	Duarte *et al* (2008)	Chemin *et al* (2010)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) General loans Jun 06 – May 07	Prosper.com (US) General loans Mar 07 Apr 07	Prosper.com (US) General loans Nov 05 – Aug 07	Prosper.com (US) General May 06 – 7 May 08	Prosper.com (US) General loans July 07 -- Dec 07	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Feb 2007 – Oct 2008	Smava.com (Germany) General data Mar07-Mar 10	Prosper.com (US) General loans June 2006	Prosper.com (US) General loans Jun06 – Jun 07	Smaca.com (Germany) General loans Mar07-Mar 10	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Jun 06- Jul 2010	Prosper.com (US) General loans May 06 – Jan 08	MyC4 (Africa – Uganda ,Kenya ,Ghana ,Ivory Coast)
C	?	-***	-***	-***	-***	-***	-***	- ***	- ***	- ***	+***	- ***	-***	+***	x
D	?	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	+***	- ***	- ***	+***	x
E	?	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	+***	- ***	- ***	+***	x
HR	?	base	base	base	base	base	base	base	base	base	+***	base	base	+***	x
Between AA – A	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Between A – B	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Between B – C	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Between C – D	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Between D – E	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Between E – HR	x	x	x	x	x	x	-**	x	x	x	x	x	x	x	x
Borrower Income	x	x	x	+*	x	x	+*	+*	+*	+*	+*	+*	+*	x	x
Borrower Income Range															
\$25,000- \$49,999	x	(-)	x	x	x	x	x	x	x	x	x	x	x	x	x
\$50,000- \$74,999	x	(-)	x	x	x	x	x	x	x	x	x	x	x	x	x
\$75,000- \$99,999	x	-**	x	x	x	x	x	x	x	x	x	x	x	x	x
\$100,000+	x	-**	x	x	x	x	x	x	x	x	x	x	x	x	x
Employment status															
Entrepreneur	?	+***	x	x	x	x	x	-***	x	x	x	x	x	x	x

Table 2-5 continues

P2P Related Studies: Effects of borrower and loan characteristics on the final interest for funded loans - OLS $I_{Final} = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Berger *et al* (2009)	Everett (2010)	Kumar (2007)	Lin *et al* (2011)	Iyer *et al* (2008)	Barasinska (2010)	Herzenstein *et al* (2008)	Herzenstein *et al* (2011)	Barasinska (2011)	Weib *et al* (2010)	Freedman *et al* (2008)	Duarte *et al* (2008)	Chemin *et al* (2010)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) General loans Jun 06 – May 07	Prosper.com (US) General loans Mar 07 Apr 07	Prosper.com (US) General loans Nov 05 – Aug 07	Prosper.com (US) General May 06 – 7 May 08	Prosper.com (US) General loans July 07 -- Dec 07	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Feb 2007 – Oct 2008	Smava.com (Germany) General data Mar07-Mar 10	Prosper.com (US) General loans June 2006	Prosper.com (US) General loans Jun06 – Jun 07	Smaca.com (Germany) General loans Mar07-Mar 10	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Jun 06- Jul 2010	Prosper.com (US) General loans May 06 – Jan 08	MyC4 (Africa – Uganda ,Kenya ,Ghana ,Ivory Coast)
Employed part time	?	(-)	x	x	x	x	x	x	x	x	x	x	x	x	x
Unemployed	?	(+)	x	x	x	x	x	x	x	x	x	x	x	x	x
Retired	?	(+)	x	x	x	x	x	x	x	x	x	x	x	x	x
Length of Employment Status	?	(+)	x	x	x	x	x	x	x	x	x	x	x	x	x
Home Ownership	?	(+)	(-)	-**	x	x	(+)	(+)	x	-**	x	(+)	-**	x	x
Picture	?	-***	-***	x	x	x	x	x	x	x	x	x	-***	x	(-)
Gender Male (base group)	base	base	x	x	x	x	x	x	x	x	- ***	x	x	base	base
Female	- ***	(+)	x	x	x	x	x	x	x	x	base	x	x	- ***	-***
Age (base group 35yrs – 60 yrs.)	Base	base	x	+**	x	x	x	x	x	x	(-)	x	x	base	x
Young(less than 35yrs)	(-)	(-)	x	base	x	x	x	x	x	x	(+)	x	x	(-)	-*
Old (more than 60)	(+)	(+)	x	?	x	x	x	x	x	x	base	x	x	(+)	x
Race White	base	base	x	x	x	x	x	x	x	x	+ ***	x	x	base	x
Black	+ ***	(+)**	x	x	x	x	x	x	x	x	base	x	x	+ ***	x
Hispanic	(+)	(+)	x	x	x	x	x	x	x	x	(+)	x	x	(+)	x
Asian	(+)	(-)	x	x	x	x	x	x	x	x	- ***	x	x	(+)	x
Beauty															
Attractiveness															
Very attractive	+*	-***	x	x	x	x	x	x	x	x	x	x	x	+*	x

Table 2-5 continues

P2P Related Studies: Effects of borrower and loan characteristics on the final interest for funded loans - OLS $I_{Final} = \alpha + \beta X + \varepsilon$	Pope *et al* (2011)	Ravini (2008)	Berger *et al* (2009)	Everett (2010)	Kumar (2007)	Lin *et al* (2011)	Iyer *et al* (2008)	Barasinska (2010)	Herzenstein *et al* (2008)	Herzenstein *et al* (2011)	Barasinska (2011)	Weib *et al* (2010)	Freedman *et al* (2008)	Duarte *et al* (2008)	Chemin *et al* (2010)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) General loans Jun 06 – May 07	Prosper.com (US) General loans Mar 07 Apr 07	Prosper.com (US) General loans Nov 05 – Aug 07	Prosper.com (US) General May 06 – 7 May 08	Prosper.com (US) General loans July 07 -- Dec 07	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Feb 2007 – Oct 2008	Smava.com (Germany) General data Mar07-Mar 10	Prosper.com (US) General loans June 2006	Prosper.com (US) General loans Jun06 – Jun 07	Smaca.com (Germany) General loans Mar07-Mar 10	Prosper.com (US) General loans Jan 07 – May 08	Prosper.com (US) General loans Jun 06- Jul 2010	Prosper.com (US) General loans May 06 – Jan 08	MyC4 (Africa – Uganda ,Kenya ,Ghana ,Ivory Coast)
Very unattractive	+**	base	x	x	x	x	x	x	x	x	x	x	x	+**	x
Happiness/smile															
Happy	(-)	(+)	x	x	x	x	x	x	x	x	x	x	x	(-)	(-)
Unhappy	(+)	base	x	x	x	x	x	x	x	x	x	x	x	(+)	base
Weight															
Somewhat overweight	+*	(-)	x	x	x	x	x	x	x	x	x	x	x	+*	x
Very overweight	(+)		x	x	x	x	x	x	x	x	x	x	x	(+)	x
Professionally dressed	+**	(+)	x	x	x	x	x	x	x	x	x	x	x	+**	x
Child with Adult in picture	+ ***	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Group Participation															
Belong to a group	x	x	-***	-**	-**	x	x	x	x	x	x	x	x	-**	-***
Group Membership fee															
Pay fee to join	x	x	-***	x	x	x	x	x	x	x	x	x	x	x	
Don't pay fee to join	x	x	-***	x	x	x	x	x	x	x	x	x	x	x	
Firm Characteristics															
Industry (2dig sic codes)															
New or existing firm															

Table 2-5 continues

Table 2-6: Analysis of factors affecting default on P2P lending websites

This table summarises the main findings in the P2P lending empirical literature on determinants driving loan default. We report the sign of the effect of the respective factors on credit allocation as positive and significant +; positive and insignificant (+); negative and significant -; variable not reported ? And variable not included x.

P2P Related Studies: Probability of default: Hazard Model of Default	Pope *et al* (2011)	Ravini (2008)	Everett (2010)	Kumar (2007)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) Empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) All loans Jun 06 – May 07	Prosper.com (US) 12 Mar 07 – 14 Apr 07 All loans	Prosper.com (US) All loans 31 May 06 – 7 May 08	Prosper.com (US) All loans July 07 -- Dec 0
Variables				
Listing Characteristics				
# listings before current	?	+*	x	x
$ amount requested	?	(+)	+***	+**
($ amount requested)^2				
# words - loan description	?	0	x	(+)
Close when funded	?	+**	+***	x
Loan Characteristics				
Final Interest rate paid	?	(+)	+***	+**
(Final Interest rate paid)^2				
Loan Duration	?	x	x	x
Business loans (relative to general loans)	+**	x	x	x
# lenders funding loan	?	x	-***	x
Loan age	x	x	+***	x
Borrower Characteristics				
Debt to Income	?	x	x	x
Bankruptcy (last 2yrs)	?	(-)	x	x
Ever Default	?	(+)	x	x
Ever default (last 7 years)	?	x	x	x
# credit lines	?	+/-	x	x
Revolving credit Balance	?	(+)	x	x

Table 2-6 continues

Table 2-6 continues

P2P Related Studies: Probability of default: Hazard Model of Default	Pope *et al* (2011)	Ravini (2008)	Everett (2010)	Kumar (2007)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) Empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) All loans Jun 06 – May 07	Prosper.com (US) All loans 12 Mar 07 – 14 Apr 07	Prosper.com (US) All loans 31 May 06 – 7 May 08	Prosper.com (US) All loans July 07 -- Dec 0
Revolving Credit amount	?	x	x	x
Credit Rating				
AA	-**	-***	x	x
A	(-)	(-)	x	x
B	(-)	(-)	x	x
C	(-)	(-)	x	x
D	(-)	(-)	x	x
E	(-)	(-)	x	(x)
HR	base	base	x	x
Borrower Income	?	x	-***	x
Borrower income range				
$25,000- $49,999	?	(-)	x	x
$50,000- $74,999	?	(-)	x	x
$75,000- $99,999	?	-***	x	x
$100,000+	?	(-)	x	x
Employment Status				
Entrepreneur	?	+/-	x	x
Employed part time	?	(-)	x	x
Unemployed	?	(-)	x	x
Retired	?	(-)	x	
Length of Employment Status		(+)	x	X
Home Ownership		(+)	(+)	(+)
Picture		(+)	x	x
Gender Male (base group)	base	base	x	x
Female	(+)	(-)	x	x
Age (35yrs – 60 yrs.)	(-)		-***	x
Young(less than 35yrs)	(-)	(-)	x	x
Old (more than 60)	base	(-)	x	x
Race Whites	base	base	x	x
Black	+***	(-)	x	x
Hispanic	(+)	(+)	x	x
Asian	(+)	(+)	x	x

Table 2-6 continues

Table 2-6 continues

P2P Related Studies: Probability of default: Hazard Model of Default	Pope *et al* (2011)	Ravini (2008)	Everett (2010)	Kumar (2007)
Data Used: Platform name (country) **Key:** Positive and sig: + Pos and insig: (+) Negative and sig: - Negative and insig: (-) Variable not included: X Variable did not survive: 0 Not reported ? New variable: (red) Empty row Significant level *** 1% ** 5% *10%	Prosper.com (US) All loans Jun 06 – May 07	All loans Prosper.com (US) 12 Mar 07 – 14 Apr 07	Prosper.com (US) All loans 31 May 06 – 7 May 08	Prosper.com (US) All loans July 07 -- Dec 0
Beauty		+***	x	x
Attractiveness		x	x	x
Very attractive	(-)	x	x	x
Very unattractive	(+)	x	x	x
Happiness			x	x
Happy	(-)	(+)	x	x
Unhappy	+*		x	x
Weight			x	x
Somewhat overweight	(+)	(+)	x	x
Very overweight	base	base	x	x
Professionally dressed	(+)	(+)	x	x
Signs of Military involvement	+*	x	x	x
Group Participation				
Belong to a group		x	+***	+*
Group Membership fee	x	x	x	x
Pay fee to join	x	x	x	+*
Don't pay fee to join	x	x	x	x
Group Rating	x	x	x	x
Group leader recommendation	x	x	x	+
Group leader reward	x	x	+***	+*
Ratio loans funded by group members	x	x	-***	
Group Size	x	x	+**	+*
Mandatory review by group leader	x	x	X	x
Group relationship	x	x	-***	x
Friend Endorsement	x	x	x	x
Borrower geographic location	x	x	-***	x
Firm Characteristics				
Industry (2dig sic codes)				
New or existing firm				

Table 2-5 continues

Table 2-7: Summary of literature and identified research gap

This table presents the summary of literature on small business venture finance from traditional off line markets, P2P lending literature on general loans and subsequently identifies a research gap in the area of P2P lending and small business loans

	Traditional Lending	Peer to Peer Lending
Small Business Loans	**Traditional lending literature on small business ventures** Three key literature streams examine: - ▪ Determinants of credit extension, with the view to understand where credit constraints kick in ▪ Determinants of cost of credit ▪ Default modelling Theory that underpins this work asymmetry of information.	**P2P lending literature on small business ventures** *Research Gap* We do not have an empirical study about the impact of P2P lending websites on the finance of small business ventures.
General Loans		**P2P lending literature on general loans** Literature Examines: - ▪ Determinants of credit extension ▪ Determinants of cost of credit ▪ Very few studies on Default ▪ Limited studies on the sustainability of the market Studies include a binary variable to compare general loans and business loans within the three streams of literature. Theory that underpins this work asymmetry of information

Chapter 3 Data and Methodology

3.0 Introduction

In this Chapter we present a description of our data and the methodology that we employ in our empirical Chapters. First, we describe the dataset and sample used in the study. Next, we put forward the econometrics methods adopted for the analysis of the data. Although different methods are employed in different empirical Chapters; on the whole Probit and Tobit models form the basis of the key econometric method adopted in data analysis. Key variables used in the study are also briefly described and finally, we conclude the Chapter by putting forward some of the limitations of the data and the adopted methods of analysis.

3.1 Data

This study is the first empirical study that looks at small business loans in the context of P2P lending. To test the hypotheses derived we have created a new and comprehensive dataset based on secondary data, extracted from the publically available electronics archives of Prosper.com (henceforth Prosper). Because Prosper is the oldest and dominant P2P lending site; it is likely to serve as a broadly useful model for examining P2P lending efforts in financing small business ventures. The data are cross-sectional in nature and they include all the information seen by potential lenders when making the lending decisions. The unit of observation is the individual loan as opposed to a firm. Because these are personal loans for business purposes, Prosper primarily underwrites them based on the owners' credit profile as opposed to the firm's credit profile.

In Figure 3-1 we show an example of a typical loan requests from Prosper while Table 3-1 summarizes all the variables collected from the loan request information. Overall there are four general types of information available in the data**.** First, the bulk of the data consists of the main credit information that Prosper obtains from the credit bureaus' reports (Experian); indicating borrower's credit history and their typical payment behavior of any previous debt obligations (including external credit scores, mortgage payments and any delinquencies or judgements). With the exception of credit scores, all other credit information is publicised on the website. Instead of the raw credit scores obtained from credit reports, Prosper assigns an internal generated credit grade to each potential borrower based on their credit score and credit history, and publishes this internally generated credit grade on the website.

...........Figure 3-1 goes around here...........

Second**,** the data contains two types of self-reported information shared by the potential borrowers on the website: (i) obligatory information which includes employment status and stated income of the potential borrowers; which is verified by Prosper and (ii) optional information in the form of a pictures and a free-form text elaborating as to why potential borrowers are good loan candidates; which is not verified by Prosper. This optional information often includes items such as intended use of the proceeds and explanation of poor credit grades. Because there are no small business specific demographic data available from the data, such as firm age, firm size, industry distribution of firms etc.; we used the optional data to create some of the demographics variables. For example, we were able to code the optional content for industry classification of the small business ventures and the firm age (classified only as a binary variable new firm or existing firm). We were able to classify the data according to loan purpose (for example working capital, capital investment etc.); and we were also able to create two additional variables (include picture and elaboration) as indicators of whether the potential borrower has availed additional information to attenuate information asymmetries.

Third, the data also contains loan specific information including loan amount, interest rates offered by the potential borrower, the interest rates paid by those who manage to get funded, as well as an indication of any previous loans attained through Prosper (a variable which we use as a proxy for reputational effects).

Finally, we have information on the lending decision outcome (loan funded or rejected). We were also able to obtain full information on the default activity of all funded loans from Prosper. The loans have a fixed maturity (36 months or 60 months) with repayments divided in equal monthly instalments. If the monthly payment is made on time, the loan status for that month is considered current. If a monthly bill is not paid, the loan status will be changed to 1 month late, 2 months late etc. If a loan is late for 60 consecutive days (but less than 90 days), it is sent to a collection agency. If a loan is late for more than 90 days, it is considered to be in default. All defaulted loans are reported to credit reporting agencies and can affect borrowers' credit scores. Borrowers who default on their loans are not permitted to borrow using Prosper again.

...........Table 3-1 goes around here...........

3.2 Sample

In total, we have a population of 28,904 loan requests, 4,046 (13%) of which were successfully funded and 24,858 (87%) of which were rejected. The analysis that follows, however, is based on a sample of 14,537 loan requests. We use a sample rather than the entire population because the short and long text descriptions from each loan request must be read and hand-coded when developing key variables. We adopted a simple random sampling technique, where every second case was randomly selected from the population. With simple random sampling, there is an equal chance that each unit from the population could be selected for inclusion in the sample.

To obtain the coded variables (industry, firm age and loan purpose), 5 postgraduate students were employed to code the sample of 14,537 loan requests (i.e. total 6 coders all sitting in 1 room for 30 days). As a starting point, approximately 10 percent of the data (1251 cases) were selected from the sample and coded by all 6 coders together to determine the unified code for each variable. Coders were then paired for cross referencing and inter-coder reliability purposes, resulting in three groups of two coding the remaining 90 percent of the data (per pair coded approximately 3750 cases). In instances where the pair of coders disagreed about classifying data – the case was brought to the attention of the team, discussed then classified. Ten percent of each group’s coded cases were checked for accuracy by a second pair of coders. On the rare occasion that coders made a large number of errors, they were asked to redo the coding and a second accuracy check of all recoded cases was performed.

In cases where there were no agreements even after discussion of cases, the case was dropped. We applied the following filtering criteria based on missing values: 766 cases (5 percent of initial sample) were dropped as they could not be classified by industry; 929 (6 percent of initial sample) were dropped as they could not be classified by firm age; 202 cases (1 percent of initial sample) were dropped as they could not be classified on the loan purpose; and 104 cases (0.7

percent of initial sample) were dropped as they had unidentifiable pictures. This resulted in a final sample of 12526 loan requests (from 7834 small firms) of which 1417 (11 percent) were funded loans. Table 3-2 shows the coder's agreement for the 1250 cases based on Feiss Kappa agreement measurement. Fleiss kappa values range from .66 to .96, which indicate substantial agreement (based on the interpretation guide offered by Landis and Koch (1977: pp13 - 17) while Table 3 shoes the correlation matrix – depicting the correlation between all variables.

...........Table 3-2 goes around here...........

...........Table 3-3 goes around here...........

3.3 Methodology

To analyse our data we use a number of bivariate and multivariate analyses. Bivariate associations were tested using t-tests and chi-square tests. For multivariate analysis we essentially relied on Probit and Tobit regressions in determining the impact of our explanatory variable of concern on the dependent variables. We now describe the broad approach used for our estimations.

3.3.1 Economic framework

The study is based mainly on two different econometric frameworks. For the first and third empirical studies, we utilize probit models explaining the determinants of credit allocation and determinants of loan default. For the second empirical study, we use Tobit models. The following subsection describes both methodological approaches in detail.

Basic probit framework

The P2P lending market has two types of participants: borrower (firm) indexed with j and lenders indexed with i. Borrowers specify the desired loan amount L_j, and interest rate I_j they are willing to pay. The desired loan amount of applicant j is funded if there are at least N lenders willing to provide funds such that $\sum_{i=1}^{M} L_i = L_j$. The willingness of the lenders to extend credit to borrower j depends largely on the borrower's probability of default p_j. Lenders do not observe

p_j; they may however infer p_j from potential borrower's observable characteristics captured in vector x_i. Likewise, the inclination of the borrower to repay the loan may be inferred from their observable characteristics. As a starting point, in the model of credit extension/default, let $\boldsymbol{Y_i}$ be the indicator that denotes which option has transpired [eg if the loan was approved = 1, else 0; and whether the loan defaulted=1 else 0]. The probit model can be expressed as follows:

$$y_i^* = \beta x_i + u_i \tag{3 -1}$$

$$y_i = \begin{cases} 1 & \text{if} \quad y_i^* > 0 \\ 0 & \text{if} \quad y_i^* \leq 0 \end{cases}$$

y_i^* is the unobserved utility of the decision maker. It is a function of a systematic component βx, where x_i is a vector of independent variables, β is a vector of parameter estimates which may change across choices (extend/reject; default/repay), u_i is a random unobserved disturbance term. Since we observe only the systematic component of utility, we cannot predict with certainty the choice of each decision maker. We can only try to assess the probability that the decision maker will choose each alternative.

The parameters of the probit model are estimated by the maximum likelihood method.
Since y_i^* is equal to $\beta x_i + u_i$, the probability that $y_i > 0$ is equal to the probability that $\beta x_i > 0$, or, equivalently, the probability that ($u_i > - \beta x_i$). Therefore, we can write the probability that y_i is equal to one as the probability that ($u_i > - \beta x_i$) such that:

$$\begin{aligned} Pr\,(y_i = \; &= Pr\,(y_i^* > 0 | \, x_i) \\ &= Pr\,(\beta x_i + u_i > 0 | \, x_i) \\ &= Pr\,(u_i > - \beta x_i) \\ &= 1 - \mathrm{N}\,(\frac{\beta\, xi}{\phi})\ \text{[integrate]} \\ &= \Phi\,(\beta x_i) \end{aligned} \tag{3-2}$$

Φ is the standard normal cumulative distribution function (we use $\phi = 1$ above; implied by the standard normal distribution). The likelihood function for this model is:

$$L = \Pi\ [\Phi(-\beta x_i)\ \Pi\ [1 - \Phi(-\beta x_i)]$$

Basic Tobit framework

In the second empirical study, the dependent variable is censored at zero because lenders will always ensure that positive return is earned from extending credit. Standard regression analysis, while feasible, assumes that the data is uncensored and therefore the format of the dependent variable suggests a standard Tobit model. The Tobit model is expressed using a limit of zero (which will be the case for our analysis) is defined with the following standard equation:

$$y_i{*} = \beta_0 + \beta_1 x_i + u \qquad (3\text{-}3)$$

$$y_i = \begin{cases} y_i{*} & \text{if } \ y_i{*} > 0 \\ 0 & \text{if } \ y_i{*} \leq 0 \end{cases}$$

Where: y_i is the dependent variable *interest rate;* x_i is a vector of independent variables; β_i is a vector of estimable parameters, and u is a normally and independently distributed error term.
The parameters of the Tobit model are estimated by the maximum likelihood method. If the interest rate is positive (i.e., for borrowers who were funded), $y_i{*}{=}y_i$, accordingly $u_i = y_i - \beta_0 - \beta_1 x_i$. As a result, the likelihood function for a funded borrower is given by the height of the density function:

$$L_i = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_i-\beta_0-\beta_1 x_i)^2}{2\sigma^2}} = \frac{1}{\sigma} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{y_i-\beta_0-\beta_1 x_i}{\sigma})^2}}_{\phi(\frac{y_i-\beta_0-\beta_1 x_i}{\sigma})} = \frac{1}{\sigma}\phi(\frac{y_i-\beta_0-\beta_1 x_i}{\sigma})$$

For loans with negative interest rates (may also typically result from borrowers who were not funded), we only know that $y^*\leq 0$. Thus, the likelihood function is the probability that $y^*\leq 0$, which is given by:

$$\begin{aligned} L_i &= P(y_i^* \leq 0) = P(\beta_0 + \beta_1 x_i + u_i \leq 0) \\ &= P(u_i \leq -(\beta_0 + \beta_1 x_i)) \\ &= P(\frac{u_i}{\sigma} \leq -\frac{\beta_0 + \beta_1 x_i}{\sigma}) \\ &= \Phi(-\frac{\beta_0 + \beta_1 x_i}{\sigma}) \\ &= 1 - \Phi(\frac{\beta_0 + \beta_1 x_i}{\sigma}) \end{aligned}$$

In sum,

$$
\begin{aligned}
&L_i = \frac{1}{\sigma}\phi(\frac{y_i - \beta_0 - \beta_1 x_i}{\sigma}) \text{ if } \mathrm{y_i} > 0 \\
&and \\
&L_i = 1 - \Phi(\frac{\beta_0 + \beta_1 x_i}{\sigma}) \quad \text{ if } \mathrm{y_i} \le 0
\end{aligned}
$$

If we let D_i be a dummy variable that takes 1 if $y_i>0$. Then, the above likelihood function can be written as.

$$L_i = \left[\frac{1}{\sigma}\phi(\frac{y_i - \beta_0 - \beta_1 x_i}{\sigma})\right]^{D_i}\left[1 - \Phi(\frac{\beta_0 + \beta_1 x_i}{\sigma})\right]^{1-D_i} \tag{3-4}$$

Basic robustness test framework

For all empirical studies with discreet dependent variables, we conduct further robustness checks of our estimations based on marginal effect estimations. Conducting marginal effects allows us to evaluate the economic impact of all factors that we have discovered to be statistically significant.

For a continuous variable: marginal effects often provide a good approximation to the amount of change in the dependent variable (y) that will be produced by a 1 unit change in the independent variable ($x_{i.}$).

Marginal Effect of $x_{i\,=}$ limit [Pr y=1|x, $x_i + \Delta$) - Pr (y=1|x, x_i)/ Δ] (3-5)

as Δ gets closer and closer to zero

For a discreet variable: With binary independent variables (x_i), marginal effects measure discreet change, that is, how do the predicted probabilities change as the binary independent variable change from 0 to 1.

$$\text{Marginal Effect of } x_{i} = \Pr(y=1|x, x_i=1) - \Pr(y=1|x, x_i=0) \qquad (3\text{-}6)$$

Furthermore, for some of the regressions we also controlled for potential sample selection bias and/or endogeniety bias by using some variant of the Heckman sample selection regressions. Sample selection bias refers to the problem where the dependent variable is only observed for a restricted, non-random sample (Heckman, 1978); as in the case of estimating determinants driving interest rates and loan default for funded loans only. In terms of the basic set up, the Heckman selection model is a two equation model. First there is the selection regression, followed by the outcome regression taking the following form:

Selection equation:

$$z_i^* = \beta_0 + \beta_1 x_i + u \qquad (3\text{-}7)$$

$$z_i = \begin{cases} 1 & \text{If } z_i^* > 0 \\ 0 & \text{If } z_i^* \leq 0 \end{cases}$$

and a basic outcome equation:

$$y_i^* = \beta_0 + \beta_1 x_i + u \qquad (3\text{-}8)$$

$$y_i = \begin{cases} \beta_1 x_i + e_i & \text{If } y_i^* > 0 \\ - & \text{If } y_i^* \leq 0 \end{cases}$$

3.3.2 Variables

In this subsection we describe the variables used in the analysis and explain how they operationalise our hypotheses. We initially discuss the independent variables. We then move on to discuss all explanatory variables and control variables (drawn from literature) which relate to the factors that drive the credit extension decision, cost of finance and default activity.

Dependent variables

The first empirical analysis (Chapter 4) is concerned with factors that drive the probability of funding in this market. Similar to previous studies (e.g. Cowling, 1997; Cole, 1998; Coleman, 2000), the dependent variable used to measure credit allocation is a decision of whether to approve credit or not (*Approve*). This decision takes a binary form; assigned to 1 for approved loan requests and 0 for rejected loan requests. Prosper is an 'all or nothing' market; in order for prospective borrowers to raise funds, their loan request must attract 100 percent of the requested loan amount. Thus, in this study, we measure funding success when loan requests raise 100 percent of the requested loan amount; partial funding was not available to potential borrowers.

The second empirical analysis (Chapter 5) is concerned with factors that drive the cost of credit in this market. The dependent variable used in the analysis is the final interest rate paid by small business owners (*Interest rate*). Following previous studies (Cowling, 1999; Burke and Hanley, 2006) the interest rate was adjusted according to the prime rate to provide the interest rate premium so as to allow for a reliable comparison between data collected at different points in time.

In the third empirical analysis (Chapter 6) in order to answer the questions related to lender returns we first analysed determinants driving default. The dependent variable used to measure default takes a binary form of 1 if a loan defaulted or 0 if the loan has been repaid (*Default*). For the purpose of the analysis, we consider default to have occurred when the principal loan amount plus calculated interest remains (partially) unpaid when the loan reached the full maturity date.

Explanatory variables

We now turn our focus to the explanatory and control variables. Following from the theory section (Chapter 2), we grouped explanatory variables into the following categories: *Owner* *attributes*, *Firm attributes* and *Information* attributes. We also control for *Loan* attributes as well as *Industry* and *Macro-economic factors*. The regressors are explained below:

Owner Attributes

We considered two borrower characteristics that are directly related to a borrower's financial strength - homeownership and credit grade. Home ownership (*Home_owner*) is an indicator variable taking value of 1 if prospective borrowers are home owners and 0 otherwise; depicted by an active mortgage loan on the borrower's credit report. In our analysis, home ownership is a measure of borrower creditworthiness; it is indicative of stability and a prior ability to access credit to obtain a mortgage. A large body of empirical studies examining small business lending has shown that borrowers' financial strength plays a significant role in their ability to obtain credit from financial institutions (Berger and Udell, 1998; Berger *et al*, 2007); hence use this variable to test for hypotheses: *H1a, H1b, and H1c*.

Credit Grade (*Credit_grade*) is also a direct indicator of a borrower's creditworthiness. It summarizes factors related to prospective borrowers' previous experience with credit, payment of bills and any failures to repay previous loan commitments. Prior research has shown that individuals' credit scores are strong predictors of their repayment likelihood for extended credit (e.g., Avery *et al*, 2004; Berger and Frame, 2007). Prosper has seven credit grade categories ranging from 1 to 7; with 1 being the lowest credit risk category and 7 being the highest credit risk category. We expect lenders to be inclined to favour borrowers with better credit grades - since poor credit grades may indicate a history of being unable to meet previous repayment obligations. In the analysis we use this variable to test for hypotheses: *H2a, H2b, and H2c.*

Delinquencies: borrowers who have delinquent obligations (*Delinquencies*) and those who have declared bankruptcy (*Judgements*) (in the past 10 years) may be viewed as less creditworthy because they have a demonstrated history of being unable to meet their previous financial obligations. Therefore, on the one hand we may expect a negative relationship between these variables and the credit allocation, cost of credit and default. However, entrepreneurial finance literature suggests that lenders may be forgiving to entrepreneurs - such that borrowers who have previously filed for bankruptcy are still able to access credit from lenders. We use this variable to test for hypotheses: *H7a, H7b, and H7c*.

Firm Attributes

Firm age (*Firm_exist*): In literature, Firm age is typically used as a proxy for the severity of information asymmetries inherent in small business ventures. Potential lenders are expected to look favourably upon loan applications from existing firms as compared to applications from borrowers starting new businesses. Existing firms are thought to be more creditworthy because they have survived the high-risk start-up period. Our data does not include the age of the firm; however borrowers do indicate whether they are looking for funds for an existing firm or a new business start-up. In our analysis Firm age is an indicator variable assigned 1 if the borrower is looking for funds for an existing firm and 0 otherwise. We use this variable to test for hypotheses: *H4a, H4b, and H4c*.

Information attributes

Repeat Loan (*Repeat_loan*): To be successful, borrowers must create a positive reputation, engendering trust within prospective lenders. Creating trust is relatively difficult in P2P lending because of the nature of the market (borrowers and lenders never meet). It is possible for borrowers to engender trust based on re-payment history. Prior loans are visible to lenders. They can be seen as a proxy for some sort of reputation that the borrower can form on Prosper, based on re-payment activity. Hence we include a variable indicating the presence and performance of a prior loan (*Repeat_loan*) and use this variable to test for hypotheses: *H3a, H3b and H3c*.

As discussed earlier, borrowers can provide (optional) text information (*Elaboration)* about themselves, their business or about their creditworthiness, and they can post pictures (*Include_picture*). Both of these variables represent strategies which borrowers may use to reduce information asymmetries. In our analysis, these variables are measured as binary indicator variables; *Include_picture* is assigned a 1 if a borrower includes an image in the loan request; we use the elaboration variable to test for *H8a, H8b, and H8c*. Similarly, *Elaboration* is assigned a 1 if prospective borrowers choose to offer a text elaboration. We use the elaboration variable to test for *H6a, H6b, and H6c*

As discussed earlier, given that hundreds of lenders assess and screen loan requests at any given time, this may serve as an indication of the level of due diligence done per loan request; which other lenders can use as a mechanism to decrease information asymmetries. We measure the number of extended pledges per loan request (*Bid_count*); to operationalise this variable; with an expectation that the more the number of extended pledges, the likely it is for the loan request to fund. We use this variable to test *H5*

Ex post default (*ex_default*): we include a variable which identifies individual loans which were, ex post, not repaid for reasons of legal default. This particular variable is our key innovation over most P2P lending studies and provides us with a pure test of amateur lenders ability to assess borrower risk for small business loans over the P2P lending context.

Control Variables

In all our empirical models we use a number of control variables to take account of additional owner specific, loan specific, macroeconomic and industry specific factors which if omitted might otherwise lead us to draw some unsafe conclusions. Specifically we include owner employment status, income range, loan size, offer interest rate, final interest rate, industry dummies, and time dummies (as a proxy for macroeconomic activity). These are all variables which have commonly been used in previous empirical studies and tend to reflect factors which might play a key role in helping the potential lenders to assess borrower risk type.

<u>Owner Attributes</u>

We include controls for borrower income (*Income_range*) and employment status (*Employment_status*); both measured as dummy variables. Traditional markets like banks value borrowers who have a constant income or who can easily get back to the employment market and earn a constant income in the event the business fails.

<u>Loan Attributes</u>

Loan Amount ($1000): We control for loan size (*Requested_amount*). This is the amount the borrowers seek to raise using Prosper. Prosper follows an ‘all or nothing’ approach. While many factors may influence credit access, cost of finance and default activity; there is a strong incentive for individuals to select realistic loan amounts. Requesting too little capital may result in project non – delivery (Hanley and Crook, 2005), and high loan size may be deemed too risky due to moral hazard effects hence less likely to succeed. We also include a polynomial variable to control for non-linear effects of the requested loan size ($Requested_amount^2$) in all our empirical models.

Offer Interest Rate (%) and Final Interest Rate (%): Offer interest rate is the maximum interest rate offered by the borrower when asking for a loan (*Offer_interest_rate*) and Final interest rate is the maximum interest rate paid by the borrowers who were extended credit (*Final_interest_rate*). In our analysis we subtract the prime rate from both variables. It is not clear what affect the offer interest rate variable will have on the credit allocation decision; likewise it is also not clear what effect the final interest rate will have on credit allocation and default. Theoretical literature suggests that borrowers that offer high interest rates may be signalling that they are high risk (Stiglitz and Weiss, 1981). However, individual lenders - by definition smaller and less professional than traditional financial institutions - may be attracted to borrowers offering high interest rates; with the view of attracting high returns. Burke and Hanley (2003; 2006) put forward the argument that interest rates may take a non-linear form. Hence, we include a polynomial variable to control for non-linear effects of both the offered interest rates ($Offer\_interest\_rate^2$) in all our credit extension empirical models and the final interest rate ($Final\_interest\_rate^2$) in all our interest rate and default empirical models.

Industry and Macroeconomic Factors

The industry variable indicates a firm's one-digit standard industrial classification to control for industry-wide differences (*Industry*) dummies in all our empirical models. We also include time dummies to account for loans asked over (*Time*) and State dummies to account for interest rates paid - given that the interest rates may vary according to States that the loan was originated (*Region*). The list of all regressors is presented in Table 3-1.

3.3.3 Estimation strategy

Overall, we adopted the ‘general-to-specific’ strategy for all empirical estimations - starting by using data on as many available variables as possible that we might expect to be relevant in determining our independent variables (guided by literature), but discarding some variables on the basis of the statistical evidence. That is, we dropped variables from the general equations which were insignificant at the 10 percent level ($p>0.10$). The aim was to derive a parsimonious model. The key benefit of parsimony in this context is that dropping irrelevant variables increases the precision or statistical significance of the estimated effects of the remaining variables on the dependant variable.

We have three empirical Chapters. We first analysed, in a multivariate setting, the factors that drive the likelihood of loan approval (Chapter 4). As reported previously, since the dependent variable is a binary outcome, we used a multivariate Probit regression model. To test the hypotheses developed we specified the following main equation:

$$Pr\ (Approval|1) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information\ Attributes + \beta_4\ Loan\ Attributes + \beta_5 Industry_i + \beta_6 Macro_i + \mu \tag{3-4}$$

Approval takes the value of 1 if the nth loan application is granted and 0 if it is rejected. Next, we analysed factors that drive the cost of credit in this market (Chapter 5). Given that it is unlikely that lenders will accept negative interest rates; we censored the dependent variable at 0; this lead to the adoption of a Tobit regression for our estimation. We considered the following equation:

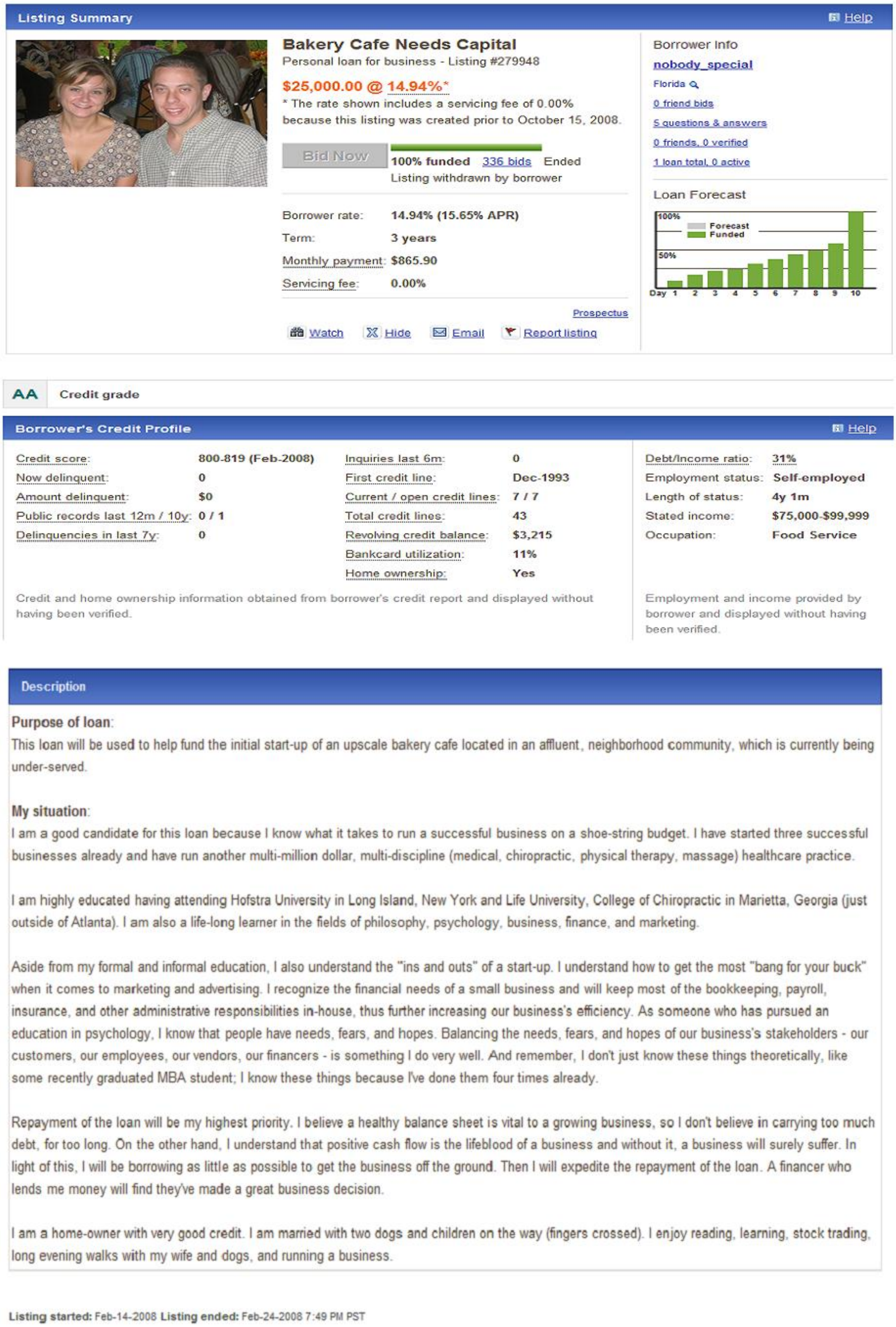
$$Interest\ rate = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ InformationAttributes_i\ \beta_4\ LoanAttributes_i + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\ Region_i + \mu \tag{3-5}$$

Finally, in line with most literature modeling default activity (Berger *et al*, 2005; DeYoung and Nigro 2005; 2008; Cowling and Mitchell, 2003), we used a binary Probit regression to test the developed predictions (Chapter 6). The dependant variable was assigned 1 if a loan defaulted and 0 otherwise. We considered the following equation:

$$Pr\ (Default) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ InformationAttributes_i\ \beta_4\ LoanAttributes_i + \beta_5 Industry_i + \beta_6 Macro_i + \mu \tag{3-6}$$

In some of the empirical studies linear and square terms of the loan size, offer interest rates and final interest rates were included to account for a non-linear relationship with dependent variables.

Figure 3-1: Example Listing of Loan request on Prosper.com



Listing Summary | Help

Bakery Cafe Needs Capital
Personal loan for business - Listing #279948

$25,000.00 @ 14.94%*
* The rate shown includes a servicing fee of 0.00% because this listing was created prior to October 15, 2008.

Bid Now | 100% funded | 336 bids | Ended
Listing withdrawn by borrower

Borrower rate:	14.94% (15.65% APR)
Term:	3 years
Monthly payment:	$865.90
Servicing fee:	0.00%

Prospectus

Watch | Hide | Email | Report listing

Borrower Info
nobody_special
Florida
0 friend bids
5 questions & answers
0 friends, 0 verified
1 loan total, 0 active

Loan Forecast
100% | Forecast | Funded | 50% | Day 1 2 3 4 5 6 7 8 9 10

AA Credit grade

Borrower's Credit Profile | Help

Credit score:	800-819 (Feb-2008)	Inquiries last 6m:	0	Debt/Income ratio:	31%
Now delinquent:	0	First credit line:	Dec-1993	Employment status:	Self-employed
Amount delinquent:	$0	Current / open credit lines:	7 / 7	Length of status:	4y 1m
Public records last 12m / 10y:	0 / 1	Total credit lines:	43	Stated income:	$75,000-$99,999
Delinquencies in last 7y:	0	Revolving credit balance:	$3,215	Occupation:	Food Service
		Bankcard utilization:	11%		
		Home ownership:	Yes		

Credit and home ownership information obtained from borrower's credit report and displayed without having been verified.

Employment and income provided by borrower and displayed without having been verified.

Description

Purpose of loan:
This loan will be used to help fund the initial start-up of an upscale bakery cafe located in an affluent, neighborhood community, which is currently being under-served.

My situation:
I am a good candidate for this loan because I know what it takes to run a successful business on a shoe-string budget. I have started three successful businesses already and have run another multi-million dollar, multi-discipline (medical, chiropractic, physical therapy, massage) healthcare practice.

I am highly educated having attending Hofstra University in Long Island, New York and Life University, College of Chiropractic in Marietta, Georgia (just outside of Atlanta). I am also a life-long learner in the fields of philosophy, psychology, business, finance, and marketing.

Aside from my formal and informal education, I also understand the "ins and outs" of a start-up. I understand how to get the most "bang for your buck" when it comes to marketing and advertising. I recognize the financial needs of a small business and will keep most of the bookkeeping, payroll, insurance, and other administrative responsibilities in-house, thus further increasing our business's efficiency. As someone who has pursued an education in psychology, I know that people have needs, fears, and hopes. Balancing the needs, fears, and hopes of our business's stakeholders - our customers, our employees, our vendors, our financers - is something I do very well. And remember, I don't just know these things theoretically, like some recently graduated MBA student; I know these things because I've done them four times already.

Repayment of the loan will be my highest priority. I believe a healthy balance sheet is vital to a growing business, so I don't believe in carrying too much debt, for too long. On the other hand, I understand that positive cash flow is the lifeblood of a business and without it, a business will surely suffer. In light of this, I will be borrowing as little as possible to get the business off the ground. Then I will expedite the repayment of the loan. A financer who lends me money will find they've made a great business decision.

I am a home-owner with very good credit. I am married with two dogs and children on the way (fingers crossed). I enjoy reading, learning, stock trading, long evening walks with my wife and dogs, and running a business.

Listing started: Feb-14-2008 **Listing ended:** Feb-24-2008 7:49 PM PST

Table 3-1: Summary of all regressors

This table list all regressors developed from our sample data from Prosper. The table gives a definition of each variable and it distinguishes between variables which were collected directly from the listing and variables which were coded.

Variable	Definition	Data Source
Home_owner	=indicator variable, taking value of 1 if borrower is homeowner, 0 otherwise;	Listing
Repeat_loan	=indicator variable, taking value of 1 if the borrower has a prior loan that has been paid off or is current at the time of the listing, 0 otherwise;	Coded
Credit_grade	=dummy indicating borrower's risk of default that takes integer values 1 = AA; 2 = A; 3 = B; 4 = C; 5 = D; 6 = E; 7= HR;	Listing
Judgements	=indicates whether the borrower declared bankruptcy within the last 10 years	Listing
Delinquencies	=number of times the borrower has been 60 or more days late with payments in the last 10 years	Listing
Employment_status	=dummy indicating the employment status of the borrower assigned to the following categories 1=fulltime;2=part-time;3=self-employed, 4=retired	Listing
Income_range	=dummy indicating income range of the borrower assigned to the following categories 1 = $0 or undefined ; 2=$25k - $49,999;3=$50k - $74,999; 4=$75k - $99,999; 5=$100k +	Listing
Existing_firm	=indicator variable, taking value 1 if firm exist, 0 otherwise;	Coded
Requested_amount	=borrower requested loan amount;	Listing
Offer_interest_rate	=interest rate offered by the borrower;	Listing
Final_interest_rate	=interest rate paid by those who were extended credit	Listing
Elaboration	=dummy indicating whether potential borrower includes a text elaboration	Coded
Include_picture	=dummy indicating whether potential borrower include a picture	Coded
Loan_term	=dummy indicating the term of the loan, taking value 1 for 36month loans and 2 for 60mnth loans	Coded
Industry	= 1 DIG SIC defined as 1=construction;2=transport and utilities;3=services;4=retail trade;5=manufacturing;6=wholesale trade;7=agriculture;8=finance and real estate	Coded
Time	=month dummies, indicating the time at which the loan request was posted on Prosper	Listing
Region	= dummies indicating all the States represented in the US	Listing

Table 3-2: The Coders' agreement measures for the three key coded variables

This table shows the coder's agreement for our three key variables.
[a] Agreement rate between the set of coders for the listings that everyone coded
[b] The interpretation of the Kappa size is based on the one suggested by Landi and Koch (1977)

Coded Variable	Coder's agreement rate[a] (%)	Fleiss Kappa	Interpretation of Kappa[b]
Firm Status			
New firm	92	.78	Substantial
Existing firm	89	.80	Substantial
Industry			
Wholesale	77	.69	Substantial
Retail	81	.73	Substantial
Services	88	.64	Substantial
Manufacturing	80	.68	Substantial
Construction	91	.71	Substantial
Mining	84	.73	Substantial
Transport	88	.64	Substantial
Agriculture	79	.66	Substantial
Finance &real estate	74	.68	Substantial
Loan purpose			
Business expansion	79	.68	Substantial
Operating expenses	96	.62	Substantial
Inventory	92	.73	Substantial
Buy existing business	96	.79	Substantial
other	79	.58	Moderate

Table 3-3: Correlation Matrix

Variable	Requested Amount	Interest rate	Income range	Credit grade	Judgements	Picture	Existing firm	Industry	Loan purpose	Homeowner	Repeat loans	Number of bids
Requested amount	1	-.17**	.18**	-.41**	-.23**	.003	-.04**	.05**	.06**	.21**	-0.16**	0.16**
Interest rate		1	-.06**	.42**	.18**	.04**	.003	-.02	-.02	-.16**	0.06**	-0.23**
Income range			1	-.19**	-.09**	.11**	-.07**	.02	.03**	.24**	0.09**	0.11**
Credit grade				1	.37**	.04**	.05**	-.05**	-.03*	-.30**	0.16**	-0.33**
Judgements					1	-.03*	.03*	-.04**	-.02	-.12**	0.10	-0.12**
Picture						1	.006	.010	-.009	.03**	-0.08	0.01**
Existing firm							1	-.020	-.12**	-.04**	0.04	0.04**
Industry								1	-.016	.04**	-0.07	0.03**
Loan purpose									1	.02*	-0.03	0.02**
Homeowner										1	0.01	0.12**
Repeat loans											1	-0.14**
Number of bids												1

Chapter 4 Factors driving small business loan approval in P2P lending

Who gets credit, who doesn’t?

4.0 Introduction

The main objective of our first empirical study is to answer the following research question: *what factors drive credit allocation for small business loan on P2P lending websites?* The basic intuition behind the research question is to determine whether P2P lending responds to known signals established in traditional lending used to predict borrower quality. If the mechanisms used by traditional lending institutions are not predictive of borrower within the P2P lending context, this may suggests that: either the signals sought by traditional lenders are ceremonial in nature and are therefore not generally predictive of entrepreneurial credit allocation outside the traditional 'offline' lending market; or that P2P lending relies on other (unique) mechanisms. Additionally, in traditional lending, many of the mechanisms available to already established firms may not necessarily be available to new business start-ups; based on the premise that new firms are the riskier offering. Hence we distinguish whether mechanisms used to cope with information issues differ between new business start-ups and existing firms in the P2P lending context.

We know from theory that some of the mechanisms which typically matter in traditional lending in terms of reducing information asymmetries include: collateral; a good credit rating; previous existing relationships with the lenders; education and greater working experience. Moreover, there is also evidence that suggests that firm age and firm size are also mechanisms that are typically used by lenders to reduce information issues.

The nature of P2P lending market is such that information asymmetries are still present. As a starting point, we assert that information asymmetries in this market may not necessarily be resolved by collateral as put forward by Bester (1985; 1987). First, in P2P lending loans are unsecured; in an event where the loans result in a default, lenders cannot seize any of the borrower's assets. Second, unlike traditional lending, if a borrower is extended credit, it is typically underwritten by a large number of individuals. If the loan results in a default, trying to gather lenders to collectively chase the borrower for the loss is unlikely; given that P2P lending can be thought of as a collection of fairly independent lenders. Third, lenders may offer as little as $50 dollars per loan request. If the loan defaults, the loss per individual lender is not too big; hence it may not be worthy to chase borrowers to try and recover the extended funds.

In contrast, if banks give a business loan, they are only dealing with one case; accordingly it is feasible to chase borrowers. Moreover, they typically are dealing with larger amounts (relative to $50 per lender) which are worth the chasing effort in an event where the loans result in default. Hence, collateral as a mechanism of reducing information asymmetries and as a signal of borrower quality is very limited in the P2P lending context; mainly because it may still be difficult to operationalise. We argue however, that the fact that borrowers indicate whether they are home owners or not in the loan requests can still provide useful information to potential lenders.

For example, if borrowers have previously managed to attain a mortgage loan (from elsewhere) and have not defaulted; this could be useful information, notifying lenders about the borrowers' creditworthiness. Following this argument, we postulate that:

H1a small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are more likely to be extended credit by prospective lenders

A related issue is that of borrower credit ratings as a means to appraisal borrower quality (Berger et *al*, 2005). Given the inability of potential borrowers to commit to pay back the loans in P2P lending (since the loans are unsecured); we argue that prospective lenders could condition the decision of whether to extend credit and the terms of credit on an individual's credit history captured by a credit score such that:

H2a small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are more likely to be extended credit by prospective lenders

Theory asserts that information asymmetries may also be resolved by established relationships between borrowers and lenders (see Sharpe, 1990; Greenbaum *et al*, 1989; Boot and Thakor, 1994; Petersen and Rajan, 1995; Berger and Udell, 1995; Cole, 1998). In traditional lending, by close and continued (physical) interaction, potential borrowers may provide the lenders with sufficient information about the firm's affairs; accumulated over time. The resulting information allows for inter-temporal arrangements, reducing credit rationing (Cole, 1998) and lowering aggregate cost of capital (Berger and Udell, 1995; Petersen and Rajan, 1995). In the P2P lending context, however, relationships in the traditional sense may not be feasible. First, borrowers and lenders never meet; hence interactions – which form an important element in reducing information asymmetries in traditional lending – are unavailable. Second, due to the sheer volume of potential lenders who could potentially extend credit per loan request; it may be impractical to form relationships as per Petersen and Rajan (1995). We contend however that borrowers can still forge reputational effects based on repayment behavior (Diamond, 1989). Information on whether potential borrowers have had a loan granted is visible to all prospective lenders; as well as the borrower's repayment behavior (conditional of the loan being granted). Borrowers that successfully pay back the loans and build some repayment history may build a good reputation. Therefore, we hypothesise that:

H3a small business borrowers who have successfully paid back a previous loan are more likely to be extended credit by lenders

It is also agreed from theory that older firms (borrowers) who are thought to have longer track records are likely to reduce information issues (Bates and Nucci, 1999; Dunne *et al*, 1989; Good and Graves, 1993; Honjo, 2000). In the context of P2P lending, borrowers are not required to indicate their age or the age of the firm. They do however indicate whether they are seeking funds for an existing business or a new business start-up. Whether firm status ultimately plays a similar role in P2P lending is unclear; especially given the fact that lenders are amateurs. However, we would expect that P2P lenders, like traditional lenders, ultimately act to rationally assess the quality of the firms such that a firm with demonstrable history or track record reduces would provide credibility even to amateurs. This logic leads to the following hypothesis:

H4a existing firms are more likely to be funded (relative to new business start-ups) in the P2Plending context

Information asymmetries in the context of P2P lending may also be solved by new mechanisms unique to this market. For example, unlike traditional lending, in P2P lending hundreds of potential lenders assess and screen the credit requests from small business venture at any given time. It could be argued that individuals understand each other's business better than traditional lenders. For example, some individuals may have knowledge about the area where the borrower wishes to start the business, others might have expertise in the product or the technology or the feasibility of the business. Thus, although prospective lenders do not physically meet the potential borrowers looking for business funds, they do see the number of other prospective lenders extending credit to a single loan request. We contend therefore that when lenders see many others extending credit to a firm, they may gauge this as an indication that extensive due diligence has been conducted by those who have already extended credit. Hence it is plausible that:

H5 the likelihood of a partially funded loan receiving additional credit will increase with the total number of already extended offers

Potential borrowers in P2P lending may include additional (optional and unverified) personal and/or financial information in their loan request in the form of text descriptions to try and attenuate information asymmetries. These descriptions often provide potential borrowers with an opportunity to explain to prospective lenders any previous failures, credit score deterioration, delinquencies, judgements or bankruptcies. The common perception in the entrepreneurial finance literature is that traditional lenders such as venture capitalists often adopt a tolerant, flexible and forgiving attitude to business failure. Studies by Zacharakis *et al* (1999) and Cope *et al* (2004) put forward that business failure is not automatically considered a 'black mark' by the venture capital community. Thus, a venture capitalist's decision to invest in a small business venture is not automatically negatively affected by a previous experience of failure. Instead, borrowers may use the interaction with venture capitalists to explain any previous failures such that failure may be considered to be part of the learning process.

In the context of P2P lending we argue that although there are no human interactions, borrowers are given an opportunity to explain to lenders any delinquencies, judgements or bankruptcies through the (optional) text elaboration such that borrowers who have previous failures may still be able to access funds on P2P lending websites. If lenders believe that the delinquencies, judgements or bankruptcies signal something relatively permanent about the potential borrower's unobservable characteristics, then it may be optimal for lenders to limit future credit. But if the circumstances surrounding previous failures are temporary (such as adverse income shock), those individuals who have just shed their previous obligations may be a good future credit risk (Chatterjee et al 2007; Behr *et al*, 2004). We argue therefore that borrowers may use the (optional) text elaboration to give account of the previous failures as well as give clarity on current repayment and bankruptcy status so as to ensure that potential lenders use this current repayment and bankruptcy status to decide whether to extend credit. Hence, it is possible that lenders in the P2P context, similar to venture capitalists, may be forgiving such that:

H6a small business borrowers who use text elaborations are more likely to be funded in the P2P lending context

One might counter the argument above however based on banking literature (Berger and Udell, 1990; Cressy, 1996). In the (unsecured) loan market, the fact that borrowers have defaulted before could actually be a big strike; a default may signal something about the borrower's future ability to repay and leads to a drop in the individual's credit score. Consequently, we hypothesise that:

H7a small business borrowers who have previous failures are less likely to access funds in the P2P lending context

Another unique feature of P2P lending is that potential borrowers may include (optional and unverified) pictures in their loan requests. These pictures often provide potential lenders with the context of the business (for example a picture of the business premises or the picture of the product). The pictures may also attempt to humanize the lending process; as it could also be an image of the potential borrower. We argue therefore that it is plausible that in the absence of human interactions, potential borrowers that include pictures could be giving prospective lenders information about the business context, or the product or even about prospective borrowers which may be used by lenders to reduce information asymmetries. This additional information may help attenuate information issues more favourably such that it increases the chances of funding success. However, it is also plausible that inclusion of these pictures may result in a negative (discriminatory) outcome. Hence we posit that:

H8a small business borrowers who post pictures are more (less) likely to get funded by lenders in the P2P lending context

The Chapter is organised as follows: we first present the descriptive statistics and simple bivariate association between key variables reflecting borrower quality and credit allocation. We then proceed with the regression analyses; reporting results for both statistical as well as the marginal effects. Finally we report results for all robustness checks conducted. We conclude the Chapter with an overall summary of our findings.

4.1 Descriptive statistics

In Table 4-1 we present the summary statistics (means and standard errors) for our explanatory and control variables. Statistics are presented separately for all loan requests (column 1); declined loans (column 2) and approved loans (column 3). In column 4, we show *t-tests* to determine whether the mean values for the funded and declined loans are statistically different.

...........Table 4-1 goes around here...........

Column 1 of Table 4-1 shows summary statistics for the sample of 12526 loan requests (from 7834 small firms); from which it is possible to profile the type of small business ventures that approach this market for funds. Using various characteristics from the loan requests, we found that typical firms approaching this market are started or owned by borrowers who are relatively in a poor credit situation: majority (76 percent) fall into Prosper's lower credit grade categories (B, C, D E HR); on average they have 6 delinquencies and at least 1 judgement record, indicating that they have previously failed to pay back loan commitments. It would appear that these owners seem to be pursuing the business venture either as a side-line to their existing works or as a hobby, given the fact that 60 percent of the sample indicate full time employment as their labour force activity. New firms make up around 30 percent of businesses approaching this market; the remaining 70 percent is made up of already established businesses. Relative to a representative sample of US small firms of which new ventures form around 10 percent, it appears new firms are over represented in the P2P context. In terms of personal wealth, just under half of the prospective borrowers own their homes.

On average, these small business borrowers are looking for small amounts of money ($10,430); they are willing to pay a high price for credit (24 percent) and the majority (55 percent) of them are looking for working capital. In terms of industry distribution, 82 percent of the sample is found in the retail, services, or finance industries. Finally, Figure 4-1 shows the regional spread of where the loan requests are typically originated – majority of the small business borrowers seeking funding are from California, Florida, Texas, New York, Atlanta, North Carolina, Illinois, and Virginia. These are regions also typical in attracting Venture Capital.

...........*Figure 4-1 goes around here*...........

Moving on to column 3 of Table 4-1, we observe that of the 12526 loan requests 1417 became funded loans; which translate to a success rate of 11 percent. So, this means that almost 9 out of 10 of those requesting a loan will not get it. In Figure 4-2, we illustrate the effort that goes into a typical loan request. We observe from this illustration that just over half of the funded loans (N = 745) were funded on the first attempt; 325 loans (23 percent) were funded after the second attempt and 347 loans (25 percent) resulted after 3 or more loan requests. These figures seem to suggest that it may be difficult to get a loan in this market.

...........*Figure 4-2 goes around here*...........

So far we have profiled the type of firm that looks for funds in this market and we have painted a picture on the effort, i.e. how easy or difficult, it may be to access funds from this market. Next we attempted to gain insight on factors associated with funding success. Hence, from Table 4-1 column 2 and column 3 we show the summary statistics of declined and funded loans respectively. In general, when comparing mean difference between funded and rejected loan requests for our explanatory variables, based on a standard *t –test*, we found statistically significant differences associated with credit approval are related to: borrower credit grades; delinquencies; judgements; labour force status; home ownership; income range; firm age, repeat loans, and inclusion of pictures. All these associations are significant at the 0.05 level or less. No statistically significant differences were found for inclusion of text elaboration. In terms of our control variables, we found that statistically significant differences associated with credit approval are loan size, interest rates and employment status; these associations were all significant at the 0.05 level. No statistically significant differences were found for income range, region of the United States and or the macroeconomic factors.

More specifically, typical business owners likely to be associated with funding success tend to be of higher credit quality (usually found in low risk credit grade categories AA). Most are homeowners (58 percent of funded loans vs. 48 percent of rejected requests indicate home ownership) and they have also previously failed to honour a loan (funded loans show an average of 3 delinquencies and 1 judgement). We further observe from Table 4-1 column 3 that new firms make up around 30 percent of firms getting funds in this market and existing firms make up 70 percent of firms getting funding, maintaining the same proportion as the firms asking for funds. This seems to suggest that lenders in this context may be indifferent about the firm status in the credit approval decision. The fact that lenders on P2P websites are lending to 30 percent of new businesses however suggests that this is a market that is favourable to new businesses; given that in a typical population of small business ventures in the US, nascent firms make up about 10 percent of the population.

Looking at our information variables, almost 40 percent of funded loans come from repeat loans, which seem to suggest an association between building some reputation in this market and access to funds. Moreover, 56 percent of loans include a picture (48 percent for declined loans) and on average it takes 191 lenders to finance a loan. Finally, in terms of loan attributes, we observe that funded business owners ask for an average of \$7,920; they receive an average of \$7,864 which seems to suggest that they get what they ask for; on average they set a maximum interest rate of 21 percent on their loan request and had an average final interest rate of 18.5 percent on funded loans.

4.2 Regression analysis

Having commented on the simple bivariate relationships between our variables in the previous section, we now move on to investigate the observed relationships in a multivariate analysis. The probit models are estimated based on equation (4 -1) using STATA V12.0. In all the estimations, we conduct a general to specific procedure (Darlington, 1990) for a set of regressors (defined in the previous chapter; found in Table 3-1) which provide various measures of owner attributes, firm attributes, information attributes as well as measures to capture loan attributes and industry conditions. We also include dummies controlling for time and macroeconomic effects.

$$Pr\ (Approval|1) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information\ Attributes + \beta_4\ Loan\ Attributes + \beta_5 Industry_i + \beta_6 Macro_i + \mu \qquad (4\text{ -}1)$$

Our estimation results are shown in Table 4-2; which is made up of four columns. The first column shows the probit results for the general specification; which includes all the regressors that could possibly drive credit allocation (identified from literature). We tested down all regressors to derive a parsimonious model shown in column 2. The testing down approach involved dropping variables from the general estimation which were insignificant at the 10 percent level (p >0.10). As stated previously, the key benefit of parsimony in this context is that dropping irrelevant variables increases the precision or statistical significance of the estimated effects of the remaining variables on the credit allocation decision.

...........Table 4-2 goes around here...........

The results of the Likelihood Ration (LR) test confirm that the parsimonious model is a better estimation; the null hypothesis that the excluded regressors collectively have no role in predicting our dependent variable is decisively accepted by the LR test (LRX^2= - 1924, df = 102, p <0.10). The third and fourth columns subsequently report the determinants of credit approval separate for existing firms and new business start-ups.

Collateral (typically proxied by homeownership) as a mechanism of reducing information asymmetries and as a signal of borrower quality is not available in the P2P lending context mainly because loans are unsecured. Therefore, information asymmetries in this market may not necessarily be resolved by collateral as put forward by theory (Bester 1985, 1987; Stiglitz and Weiss, 1981; and Besanko and Thakor, 1994). Consequently, collateral in the context of P2P lending becomes less important. Our results show however that given the fact that borrowers indicate whether they are home owners or not in the loan requests provide useful information to potential lenders. We show, in column 2 of Table 4-2, that similar to traditional lending, the supply of loans flows to the least risky entrepreneurs who are homeowners (H1a); confirming the importance of homeownership (as useful mechanisms of eradicating information asymmetries in the P2P lending context (Berger and Frame, 2007). Not as a form of collateral, but in the form of borrower reputation as stipulated by Diamond (1989). Borrowers who are consistent in mortgage repayments seem to build a positive reputation, thus gaining access to loans in this context.

Similarly, the supply of loans flows to less risky entrepreneurs with high credit ratings (H2a) and to those indicating that they are repeat borrowers (H3a) with previous established repayment history within the P2P lending context. In fact our proxy of previously existing relationship Repeat_loans does not appear in the probit results as shown in Table 4-2 because it predicts funding success perfectly. In traditional lending, by close and continued (physical) interaction, potential borrowers may provide the lenders with sufficient information about the firm's affairs; accumulated over time. The resulting information allows for inter-temporal arrangements, reducing credit rationing (Cole, 1998). In the P2P lending context, however, relationships in the traditional sense may not be feasible. First, borrowers and lenders never meet; hence interactions – which form an important element in reducing information asymmetries in traditional lending – are unavailable. Second, due to the sheer volume of potential lenders who could potentially extend credit per loan request; it may be impractical to form relationships as per Petersen and Rajan (1995). Although there are no physical interactions between borrowers and lenders in this context, our results corroborates Diamond's (1989) findings regarding the importance of building a reputation with a prospective lender; implying that a presence of a track record and the building of a reputation matters in reducing information asymmetries and adverse selection issues in the P2P lending market see Sharpe, 1990; Greenbaum et al, 1989; Boot and Thakor, 1994; Petersen and Rajan, 1995; Berger and Udell, 1995; Cole, 1998).

Our results also demonstrate that firm level characteristics have little impact on loan supply. We observe from column 2 of Table 4-2 that our variable Existing_Firm is not significant. Hence we do not find support for H4a. This result is counter to what is typically seen in banking literature (see Cole, 1998; Coleman, 2000); where the age of the firm (entrepreneur) is an important determinant of access to finance. This is largely due to the fact that older firms (entrepreneurs) are thought to have longer track records; hence they are likely to reduce information issues (Bates and Nucci, 1999; Dunne et al, 1989; Good and Graves, 1993; Honjo, 2000). One plausible explanation for this observation is that although lenders in this context act like debt financiers, they appraise risk more like the decision-making of equity funders who focus more on people rather than the business itself. Another plausible explanation for this observation is that lenders may be extending credit based on personal idiosyncrasies – for example they may be extending credit based on philanthropy (Agarwal et al, 2011; Mollick, 2013); simply because they identify with the course of the entrepreneur, regardless of firm age.

Interestingly, our results show support for Stiglitz and Weiss (1981) assertion that reducing information gaps in the form of conducting due diligence improves access to loans. We observe from column 2 of Table 4-2 that information asymmetries in the context of P2P lending may also be solved by the ‘crowds’. For example, unlike traditional lending, in P2P lending hundreds of potential lenders assess and screen the credit requests from small business venture at any given time. It could be argued that individuals understand each other’s business better than traditional lenders. For example, some individuals may have knowledge about the area where the borrower wishes to start the business, others might have expertise in the product or the technology or the feasibility of the business. Thus, although prospective lenders do not physically meet the potential borrowers looking for business funds, these lenders do see the number of other prospective lenders extending credit to a single loan request. Our results show that ‘crowds’ somehow help to reducing information asymmetries (H5); the variable Bid_count is positive and significant at the 0.01 level; which suggests that prospective lenders perceive loan requests attracting a large number of potential lenders to have conducted a great deal of due diligence. It is important to highlight however that although our results support the model of Stiglitz and Weiss (1981); P2P lending differs in that unlike traditional banks where due diligence is conducted by an individual, due diligence is conducted by the crowd – a new unique mechanism of reducing information asymmetries and adverse selection risk in P2P lending.

The common perception in the entrepreneurial finance literature is that traditional lenders such as venture capitalists often adopt a tolerant, flexible and forgiving attitude to business failure. Studies by Zacharakis et al (1999) and Cope et al (2004) put forward that business failure is not automatically considered a ‘black mark’ by the venture capital community. Thus, a venture capitalists decision to invest in a small business venture is not automatically negatively affected by a previous experience of failure. Instead, borrowers may use the interaction with venture capitalists to explain any previous failures such that failure may be considered to be part of the learning process. In the context of P2P lending our results show that although there are no human interactions, borrowers are given an opportunity to explain to lenders any delinquencies, judgements or bankruptcies through the (optional) text elaboration such that borrowers who have previous failures may still be able to access funds on P2P lending websites. If lenders believe that the delinquencies, judgements or bankruptcies signal something relatively permanent about the potential borrower's unobservable characteristics, then it may be optimal for lenders to limit future credit. But if the circumstances surrounding previous failures are temporary (such as adverse income shock), those individuals who have just shed their previous obligations may be a good future credit risk (Chatterjee et al 2007; Behr et al, 2004). Interestingly, however it seems the stories which these borrowers tell do not influence funding success (H6a); our information variable Elaboration is not significant.

However, our results show that lenders respond positively to pictures such that small business borrowers who include them in their loan requests are more likely to be extended credit (H8a); our variable Include_picture is positive and significant at the 0.01 level; suggesting perhaps that inclusion of a picture somehow humanizes the process. Hence our results seem to suggest that in the absence of human interactions, potential borrowers that include pictures could be giving prospective lenders information about the business context, or the product or even about prospective borrowers which may be used by lenders to reduce information asymmetries. This additional information may help attenuate information issues more favourably such that it increases the chances of funding success. Pope and Sydnor (2013), give an alternative explanation for this observation – asserting that lenders may be reacting to pictures based on beauty of those in the pictures – such that more beautiful individuals are more likely to get extended credit.

Looking at the control variables, we observe that the larger the amount requested by the (prospective) small business owners , the higher the probability that their loan request will be declined, as evidenced in the negative sign and high significant coefficient of the variable Requested_amount. This seems counter intuitive to previous studies (see Hanley and Girma, 2006); where it is possible that lenders may be likely to reject lower credit requests if these projects are synonymous with underinvestment and hence pose a higher failure risk if the business is under. Interestingly, looking at the linear and square requested amount variables together, our results propose a U shape – suggesting that borrowers seeking smaller loans on the one extreme and those seeking large loans on the other extreme are less likely to be funded. In an attempt to confirm this observed relationship, Figure 4-3 plots the variable Requested_amount2.

As identified in the figure, the relationship appears to support non-linear (at a decreasing rate); rendering the effects of square term somewhat important (this was formally tested later when conducting marginal effects). Our results also report that once borrower risk is accounted for - borrowers that offer higher interest rate are more likely to be extended credit; supporting results put forward by Hanley and Girma, (2006). Finally, we find that similar to banks, prospective borrowers who indicate that they have some form of income, and those who are in full time employment are more likely to be extended credit (all these factors are significant at the 0.05 level or below).

...........Figure 4-3 goes around here...........

Next, in order to confirm the result as shown by the indicator variable, *Existing_firm*, in column 3 and column 4 of Table 4-2 we extend the analysis to determine factors associated with the likelihood of default separately for existing firms and new business start-ups as defined by equations (4-2) and (4-3).

$$Pr\ (Approval{:}\ Existing\_firms|1) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information\ Attributes + \beta_4\ Loan\ Attributes + \beta_5 Industry_i + \beta_6 Macro_i + \mu \qquad (4\text{-}2)$$

$$Pr\ (Approval{:}\ New\_firms|1) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information\ Attributes + \beta_4\ Loan\ Attributes + \beta_5 Industry_i + \beta_6 Macro_i + \mu \qquad (4\text{-}3)$$

Our estimation results are shown in Table 4-2 column 2 and column3 for existing firms and new firms respectively. The estimated probabilities of credit approval for existing firms and new business start-ups are virtually identical. In both estimations the likelihood of credit approval is determined by home ownership, credit grade, previous judgements; firm age, repeat loans, the posting of pictures, bid count, borrower income; and employment status (all these factors are significant at the 0.05 level or below).

With the exception of homeownership variables, all the other attributes determining funding success for new firms and existing firms are similar. One explanation for the modest role for home ownership in predicting credit approval for existing firms could be that much of the information may already be contained in other variables - notably the *Credit_grade* variable and the credit history variables (*Judgements* and *Delinquencies*) because credit scores generally incorporate borrower’s payment history on any credit extended to individuals.

4.3 Marginal Effects

In the preceding section, we found that typical tools used in traditional lending such as credit history, credit scores, homeownership significantly influenced the probability of loan approval; whilst firm level attributes seem less important. We also observed that mechanisms unique to the P2P lending context such as crowd intelligence and pictures also help attenuate information asymmetries. In this section of the paper, we provide a quantitative assessment of the influence of the observed results. We measure the marginal effects of all significant observations.

4.3.1 Average marginal effects

In Table 4-3 we report the estimated marginal effects of the individual regressors on the credit approval decision. By setting most regressors at their sample means, we attempt to discover the impact of changes in one regressor for an otherwise 'typical' case. The target was to demonstrate how entrepreneurs possessing different attributes exhibit credit approval probabilities. Table 4-3 has three columns. The first column traces marginal effects of a particular regressor on the likelihood of credit approval for all loans. The second and third columns subsequently trace the marginal effects separate for existing firms and new business start-ups.

...........Table 4-3 goes around here...........

We see in column 1 of Table 4-3 that the borrower quality dummy, '*Credit_grade'*, is the single most important variable in the credit approval decision. We estimate that in comparison to a small business owner with credit grade in category AA, borrowers in category A are 4 percentage points less likely to get funded. Compared to the average probability of funding, 11 percent, this translates to approximately 40 percent reduction in the likelihood of the loan application being approved. As the credit grades deteriorate, the impact becomes even more drastic.

For instance, borrowers in category C and those in category HR are 10 and 20 percentage points less likely to get funded which translate to 90 percent and 180 percent reduction in funding success respectively. Similarly, in terms of borrower quality, home ownership is also important, albeit at a lesser impact relative to credit rating. Loan applications from borrowers who are homeowners are 1.4 percentage points more likely to be funded; this translates to approximately 13 percent increase in the likelihood of the loan application being approved relative to the average probability of funding.

Moreover, our results show evidence that prospective lenders are cognisant of previous handling of credit issues when sanctioning a loan to a would-be borrower. Prospective borrowers who have had previous borrowings rescheduled due to an inability to meet repayments or who have been insolvent in the past, *Delinquencies*, are 16 percent less likely to have their applications approved. It is plausible however that the credit grade variable may be capturing much of the information already reported in the homeownership and credit history variables (*Delinquencies* and *Judgements*) because , as previously stated, credit scores generally incorporate borrower's payment history on any credit extended to individuals. This directly includes mortgage payments, such that paying the mortgage on time strengthens the credit score while late payments affect the credit score adversely.

In terms of the information variables - *Include_picture* and *Bid_count*- the impact of pictures is evident in the fact that compared to borrowers who opt not to include pictures in their loan application; we estimate that borrowers who do are 2.4 percentage points more likely to receive funding. Compared to the average probability of funding, 11 percent, this represents a 23 percent increase in the likelihood of receiving funding. It is somewhat surprising that we find evidence that pictures seem to attenuate information asymmetries between lenders and borrowers, because pictures are optional and not verified by Prosper. A natural expectation therefore would be that lenders in this market would respond little to this type of 'cheap talk'. Yet the fact that borrowers include a wide variety of (non-standardized) pictures and the market responds to them, suggest that the information on the contrary is not treated as cheap talk by lenders in this market. Perhaps the pictures do indeed serve as some element of humanizing the lending process. Interestingly, we also find strong support for the number of lenders and collective due diligence argument. For every one additional lender extending credit to a loan request, this increases the funding success of the loan by 0.6 percent; for every 10 persons, the funding increases by 6 percent, and for every 100 people, the likelihood of funding success increases by 60 percent.

When we examined the marginal effects of control variables, we found that dropping the requested loan by $1,000 in the amount requested significantly increases the likelihood of getting funds by 23 percent. Interestingly, dropping the requested loan amount by $1,000 seems to have the same effect on the chance of being funded as including a picture in the loan request. As identified earlier both the requested amount and squared requested amount obtained significant coefficients, suggesting a nonlinear relationship between loan size and the likelihood of funding. We see from column 1 of Table 4-3 that the marginal effect of the square variables is small but significant (0.9 percent). Interestingly, once we control for credit risk - borrowers that offer a 1 unit increase in interest rate, increase their likelihood of funding success by 0.6 percent. Increasing the interest rates by a factor of 10 results in the likelihood of funding success increasing my 6 percent - suggesting that borrowers really need to dig need and offer higher interest rates. Finally, in terms of labour force participation, being evasive about the labour force activity does not bode well with potential lenders – it decreases the likelihood of funding by 4 percentage points. This translates to a 36 percent decrease in the likelihood of funding; a much tougher punishment than that received than those who state that they are self-employed.

4.3.2. Beyond average marginal effects

The results presented thus far measure marginal effects. The biggest problem with this approach however is the fact that it only produces a single estimate of the marginal effect. No matter how 'average' is defined, averages can obscure difference in effects across cases. It is of course possible that the effect of certain variables will differ across the distribution of small business ventures; especially given that small business ventures by nature are heterogeneous. Moreover, the interpretation of the concept of 'average' may be rendered difficult by the discrete nature of many of our regressors. Therefore in this section we conduct marginal effects at representative values; which could be both intuitively meaningful, while also showing how the effects of variables vary by other characteristics of the small business borrowers.

Table 4-4 and Table 4-5 confirms that moving beyond the average effect of each variable on credit approval uncovers a large amount of heterogeneity in credit approval. As a starting point, in Table 4-4 we show the estimates of how the marginal effects differ over a range of different requested loan sizes across different borrower credit grades (all other variables were kept at their mean). What we observe from the Table 4-4 suggest that credit approval does decrease as the loan size increases (across all credit grades); larger loans are more risky. As identified previously, we see from the table the relationship between loan size and credit approval does appear to be linear. Also, the effect of credit grades across different loan sizes is immediately eminent.

...........Table 4-4 goes around here...........

In Table 4-5 we take the analysis a step further by compiling predicted probabilities for credit allocation at representative values of owner attributes, firm attributes, loan attributes and information attributes. We selected cases from our data file - a mix of extremes (highest and lowest credit risk) and a few cases in-between to illustrate our point (of moving beyond marginal effects). The baseline case is an individual with *credit grade C, full time employment, income range $25k - $49,999k, include a picture, with delinquencies, who owns their home* – this borrower has a 7 percent likelihood of funding success. Compared to the baseline case, a borrower with *premium credit grade AA, full time employed income range $50k - $74,999k, no past due loans, who is a homeowner*, has a 20 percent chance of getting a loan request funded; this translates to a probability of funding that is almost 3 times more likely than the base case. Whilst, compared to the baseline case, a high risk borrower with *HR credit grade, self-employed, income range $1k - $24,999k with delinquencies and who rents their home*, has 1 percent chance of being funded.

...........Table 4-5 goes around here...........

Interestingly, we observe from the table that for our extreme cases (AA and HR) - the pictures certainly have a bigger impact for those in premium credit grades; whilst they seem to be doing very little for those who already have bad credit ratings. For premium borrowers not including a picture (for otherwise similar borrowers) reduce their funding success by 3 percentage points. In comparison, the higher risk borrowers including a picture still keeps the probability of funding success in lower single digits, which seems to suggest that for this cohort including a picture does very little to attenuate information asymmetries.

What's more, although on average borrowers who have failed to honour their credit obligations in the past - with delinquencies and judgements - are less likely to be funded. Our results seem to suggest that for borrowers who do bounce back and manage to rectify their previous credit discretions and work up their credit rating, the fact that they have previously failed to honour their commitment is overlooked. They are able to attract capital. For instance Table 4-5 shows that a borrower with the profile*: Full-time employed, credit grade A, income range $25k - $49,999k, past due loans and judgements, homeowner, existing firm, image, elaboration* has a 15 percent chance of funding success. They actually outperform premium borrower *Self-employed, premium credit grade AA, income range $25k - $49,999k, no past due loans and no judgements, rent, new firm, no image, elaboration* by 3 percentage points. This observation seems to support the notion that, perhaps similar to venture capitalists, lenders on Prosper are somewhat forgiving.

4.4 Robustness Check

In the preceding sections, we found that credit grades significantly impacts the probability of loan approval. Owner, information and loan variables were a little less important, while firm variables were not important. In this subsection, we conduct a number of additional analyses to examine the robustness of our results.

4.4.1 Removing the credit grade variable

As a starting point, we looked at the robustness of our results based on the credit grade variable, since this is the single most important determinant of loan approval. As stated earlier, it is plausible that the credit grades may be capturing much of the information already reported in other variables – notably the *Home_ownership* variable and the credit history variables (*Judgements* and *Delinquencies*) because credit scores generally incorporate borrower's payment history on any credit extended to individuals. To see if this is indeed the case, we re-estimated the parsimonious models (defined in Table 4-2 columns 2, 3 and 4) without the credit grade variable; results of which are shown in Table 4-6.

...........Table 4-6 goes around here...........

In this analysis, noticeably the *Home_ownership* and *Judgement* variables which were jumping in and out of significance in the previous model specifications - are now statistically significant for all estimations. What's more, we observe from the marginal effects on these variables that the impact of these variables on the likelihood of funding has also strengthened. In the end credit grades, home ownership, delinquencies and judgements all measure the level of trust. Since lenders do not have personal relationships with borrowers, they need an objective measurement which all these variables provide. The fact that being a homeowner can build trust (the same argument applies to borrowers with no delinquencies and no judgements) which may be easily translatable for an otherwise (less sophisticated) individual lender in the P2P context; which justifies keeping all these variables in the models.

Most importantly, we observe that the estimated probabilities of loan approval for our key information variables are generally consistent with our previous findings; regardless of exclusion of the information on credit grades including of a picture increases funding success; and the. What's more, firm level characteristics are still insignificant, consistent with our previous results. Finally, our results also confirm that coefficient signs on the control variables were also generally consistent.

4.4.2 Changes in the specification of the dependent variable

We recall that Prosper is an 'all or nothing' market; small business borrowers need to raise hundred percent of the requested amount in order for the loan to be funded. Although this is the case, Prosper provides information on the fraction funded for each loan request. We use this measure as an alternative dependent variable to provide a robustness check on factors driving credit approval. The empirical specification is the following Tobit regression:

$$\text{Fraction\_funded} = \alpha + \beta_1\, Owner_i + \beta_2\, Firm_i + \beta_3\, Information + \beta_4\, LoanAttributes_i + \beta_5 Industry_i + \beta_6 Macro_i + \mu \qquad (4\text{-}3)$$

Where the independent variables are the same as those defined in Table 3-1 (Chapter 3). All in all, our results, shown in Table 4-7, confirm our earlier findings; the estimated coefficients on all our explanatory variables retain very similar coefficients in terms of statistical significance.

..........Table 4-7 goes around here..........

4.5. Chapter summary

Recalling our research question, this study considers the nature of the signals adopted by P2P lenders in attenuating information asymmetry problems. These include mechanisms typically used in traditional lending such as home ownership, credit grades as indicators of borrower creditworthiness and those relating to measures of the small business. The study also considers new mechanisms adopted such as the intelligence of the crowd and inclusion of pictures – which may attenuate information asymmetries.

Overall, these results suggests the following: P2P lending depicts a new venture loan market where previously underserved early stage ventures and those looking for small amounts are able to access credit, with the relaxation of typical collateral requirements since these loans are unsecured. However, although P2P lending is a new innovation, mechanisms of eradicating information asymmetry challenges adopted in traditional finance are still valued on P2P websites. From this study, we put forward three key conclusions in terms of supply of loans: (i) our results show that the supply of loans flows to the less risky entrepreneurs (homeowners, those with high credit ratings, and repeat borrowers); highlighting the importance of reputation in this context (ii) our result show that firm level characteristics (including firm age) have little impact on the supply of loans; suggesting that idiosyncratic features like philanthropy might potentially also be at play (iii) reducing information gaps through crowds and pictures improves access to loans; where new crowds improve the conducting of due diligence, whilst inclusion of pictures help humanize the lending process. These findings are both interesting and important in that they suggest P2P lending is a low risk form of debt finance for investors; in this sense they act like commercial debt financiers. But the way they appraise risk funding opportunities is more like the decision-making of equity funders who focus on people rather than the business itself.

Our results have key implications to theory. To recap, Stiglitz and Weiss put forward four key parameters underpinning their theory of information asymmetry and adverse selection: collateral, conducting due diligence, refraining from high interest rates to avoid moral hazard and adverse selection issues, and the inferred face to face borrower-lender interactions. From our results, we find that some weights of these parameter values are likely to change in Stiglitz-Weiss model.

First, the general insight we get from is that borrower reputation, stipulated by credit grades, is the single most important determinant of credit allocation .The significance of using the credit grade helps to reduce the problem of adverse selection. The cost of defaulting will result in poorer scores – which are quantifiable to 24 percent in reduction of probability of funding. Moreover, with the advancement of internet, reputations which were previously limited within the 1-to-1 lending typology from banks, where if a borrower defaults on credit in one region or country for example, would not have an effect if the borrower were to move to another country has since have completely changed. But with internet age loan default may quickly go viral. The consequence of 1-to- many borrower-lender interactions relative to reputation over the internet and the ease with which default may go viral makes reputation to be a very important aspect within the P2P lending context.

Second, we see from our results that collateral, which was such an important determinant in reducing adverse selection issues in Stiglitz-Weiss theory, in the P2P lending it is unimportant.

Third, the general insight we get is that due diligence, although still an important factor, in the P2P lending context is conducted by the ‘crowd’. This new feature, unique to P2P lending was not taken into consideration in the Stiglitz and Weiss framework – where credit risk appraisal was done by relatively one person. Consequently we introduce collecting intelligence as a means of eradicating information asymmetry and adverse selection issues. Furthermore, our results shift focus from 1-to-1 physical interactions between borrowers and lenders inferred in Stiglitz-Weiss theory and highlights the 1-to many borrower lender typology over the internet. Effectively rendering physical contact, which was previously seen an important aspect of reducing information asymmetry and adverse selection issues in theory, relatively less important.

Furthermore, In terms of due diligence, the crowd also introduces another distinctive change based in the notion that lenders in this context may have philanthropy ambitions which they may consider when appraising credit risk. For example, a lender taking into consideration philanthropic ambitions may look for different credit risk when compared to a lender whose sole ambition is to maximise returns. This may effectively alter access to credit and subsequently the P2P borrower pool.

We contend that the lessons we leant from our results about asymmetric information and adverse selection issues are quiet different to those developed in Stiglitz-Weiss model. Reputation is very important in this market (this was not highlighted by Stiglitz). We learn that physical contact and collateral becomes less important in reducing information asymmetries. We also learn about the importance of three key new features: collective intelligence of the crowd, an aspect of philanthropy present when appraising credit risk and the general element of fun which may drive lenders when choosing to allocate credit. We contend that Stiglitz-Weiss model will have to be updated to take into consideration the facts raised above in their theory of information asymmetry in order to reflect our finding.

Table 4-1: Descriptive statistics

In this table column (1) shows descriptive statistics all funded loans. In column (2) we show statistics for all loans which were paid back whilst column (3) shows all the loans that resulted in default. Finally, column (4) presents *t-test/x^2 statistics* for differences in the means of the repaid and defaulted loans. *, **, *** stand for 0.1, 0.05, and 0.01 significance levels respectively.

Variable	(1) All loans	(2) Declined loans	(3) Funded Loans	(4) *t- test*
Number of loans	12,526	11,109	1,417	
Owner attributes				
Home_ownership	0.49	0.48	0.58	6.8***
Credit_grade				
AA	0.11	0.09	0.22	15.1***
A	0.12	0.11	0.18	7.9***
B	0.15	0.15	0.18	3.3***
C	0.18	0.18	0.14	-3.7***
D	0.15	0.15	0.15	-.07
E	0.09	0.10	0.07	-2.7**
HR	0.20	0.22	0.05	-15.5***
Delinquencies	6.1	6.5	2.9	-9.6***
Judgements	1.7	1.8	0.4	12.1***
Income_range				
$0 or undefined	0.13	0.13	0.06	-8.4***
$1 - $24,999	0.10	0.11	0.08	-3.2***
$25k - $49,999	0.27	0.27	0.28	1.4
$50k - $74,999	0.23	0.22	0.27	4.2
$75k - $99,999	0.11	0.11	0.14	3.2***
$100k +	0.16	0.16	0.17	0.83
Employment_Status				
Full time	0.60	0.57	0.75	-12.7***
Part time	0.02	0.02	0.01	2.3***
Self-employed	0.35	0.36	0.21	11.3***
other	0.03	0.05	0.03	1.4*
Firm Attributes				
Existing_firm	0.70	0.70	0.71	-2.9*
Industry				
construction	0.02	0.02	0.01	-2.4**
transport and utilities	0.02	0.02	0.02	-1.3
services	0.42	0.42	0.40	-1.2**
retail trade	0.31	0.31	0.30	1.1
finance &real estate	0.16	0.17	0.20	3.8***
agriculture	0.01	0.01	0.02	3.1
wholesale trade	0.01	0.004	0.01	2.5
manufacturing	0.05	0.05	0.05	-0.7
Information attributes				
Include_picture	0.56	0.55	0.56	0.68**
Elaboration	0.97	0.97	0.97	-2.5
Repeat_loan	0.04	N/A	0.39	N/A
Bid_count	39	20	191	1.6**
Loan Contract Attributes				
Requested amount	$10,430	$10,751	$7,920	13.4***
Funded amount	$1,821	-	$7,864	N/A
Offer_ interest_ rate	24.1	24.5	21.3	-12.1***
Final_ interest_rate	23.8	24.4	18.5	-20.1***

Table 4-2: Probit estimates of factors driving credit approval

This table reports the Probit regression results for factors driving credit approval on Prosper.com. The first two regressions present estimates for the general and parsimonious specifications for all loan requests. The last two regressions present estimates for parsimonious specifications for loans from existing firms only and for loans from new firms only. In all regressions, the dependent variable is binary taking the form 1 if credit the loan request was funded and 0 otherwise. The explanatory variables include owner, firm, and information attributes: credit grade, home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan.

The controls are requested loan size, offer interest rate, employment status, and income. Time and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Pr\ (Approval|1) = \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributes_i + \beta_5 Industry_i + \beta 6 Macroi + \mu$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively.

Variable	(1) General model (all)	(2) Parsimonious Model (all)	(3) Existing firms (only)	(4) New firms (only)
Constant	-0.332	-0.122	0.956***	0.423
	(-0.350)	(-0.138)	(2.609)	(0.566)
Owner Attributes				
Homeowner	0.134***	0.136***	0.087	0.299***
	(2.685)	(2.732)	(1.483)	(2.850)
Credit_grade (AA)				
A	-0.263***	-0.245***	-0.210**	-0.421**
	(-2.858)	(-2.704)	(-2.033)	(-1.982)
B	-0.397***	-0.369***	-0.339***	-0.498**
	(-4.019)	(-3.899)	(-3.126)	(-2.334)
C	-0.662***	-0.625***	-0.717***	-0.415*
	(-6.079)	(-6.059)	(-5.930)	(-1.863)
D	-0.921***	-0.881***	-0.906***	-0.881***
	(-7.468)	(-7.505)	(-6.636)	(-3.424)
E	-1.373***	-1.347***	-1.421***	-1.255***
	(-9.741)	(-9.748)	(-8.702)	(-4.336)
HR	-1.849***	-1.835***	-1.871***	-1.880***
	(-12.552)	(-12.639)	(-11.033)	(-6.028)
Delinquencies	-0.170***	-0.175***	-0.124**	-0.263**
	(-3.239)	(-3.342)	(-1.998)	(-2.417)
Judgements	-0.223			
	(-1.514)			
Income_range ($ 0 – unable to verify)				
$1 - $24,999	0.184*	0.214*	0.098*	0.704**
	(1.666)	(1.829)	(0.719)	(2.445)
$25,000 - $49,999	0.277***	0.301***	0.210*	0.764***
	(2.874)	(2.975)	(1.879)	(2.787)
$50,000 - $74,999	0.322***	0.344***	0.265**	0.794***
	(3.280)	(3.350)	(2.351)	(2.856)
$75,000 - $99,999	0.232**	0.253**	0.093*	0.950***
	(2.093)	(2.201)	(0.730)	(3.135)
$100 000 plus	0.106*	0.120	0.013	0.671*
	(0.957)	(1.052)	(0.106)	(2.129)
Employment_status (full time)				
Part-time	-0.450**	-0.443**	-0.492**	-0.465*
	(-2.435)	(-2.393)	(-1.989)	(-1.472)
Self-employed	-0.107*	-0.100*	-0.137**	-0.140*
	(-1.947)	(-1.820)	(-2.248)	(1.031)

Table 4-2 continue

Variable	(1) General model (all)	(2) Parsimonious Model (all)	(3) Existing firms (only)	(4) New firms (only)
Firm Attributes				
Existing_firm	0.054			
	(1.041)			
Information Attributes				
Include_picture	0.231***	0.238***	0.179***	0.435***
	(4.138)	(4.261)	(2.695)	(3.788)
Elaboration	0.140		0.511	-0.348
	(0.561)		(1.365)	(-0.795)
Bid_count	0.008***	0.008***	0.007***	0.011***
	(30.644)	(31.048)	(26.262)	(16.183)
Loan Attributes				
Requested_amount($1000)	-0.249***	-0.247***	-0.225***	-0.335***
	(-16.221)	(-16.251)	(-12.844)	(-9.675)
SQ Requested_amount	0.003***	0.003***	0.002***	0.004***
	(5.101)	(5.061)	(3.477)	(3.344)
Offer_interest_rate (%)	0.030*	0.013***	0.014***	0.010**
	(1.702)	(3.694)	(3.494)	(1.304)
SQ Offer_interest_rate	-0.000			
	(-0.982)			
Time fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Number of Observations	10,278	10,281	7,249	2,983
Pseudo R^2	0.481	0.481	0.473	0.549
Log Likelihood	-1910	-1924	-1400	-459.2

Table 4-3: Marginal effects after Probit regression

This table presents the marginal effects after Probit models based on all variables set at their means. The marginal effects for categorical variables show how Pr (Approval = 1) is predicted to change as a particular factor variable changes from 0 to 1, holding all other independent variables at zero. The marginal effect of a continuous variable measures the instantaneous rate of change, which may or may not be close to the effect on Pr (Approval=1) of a one unit increase in the independent variable. *, **, *** stand for 0.1, 0.05, and 0.01 significance levels respectively.

	All firms	Existing firms only	New firms only
Owner attributes			
Homeownership	0.014	0.009	0.025
	(0.005)***	(0.006)	(0.009)***
Credit_grade (ref AA)			
A	-0.043	-0.022	-0.036
	(0.016)***	(0.011)**	(0.018)**
B	-0.063	-0.036	-0.042
	(0.017)***	(0.011)***	(0.018)**
C	-0.099	-0.076	-0.035
	(0.018)***	(0.013)***	(0.019)*
D	-0.131	-0.096	-0.075
	(0.019)***	(0.014)***	(0.022)***
E	-0.173	-0.151	-0.107
	(0.020)***	(0.017)***	(0.024)***
HR	-0.201	-0.199	-0.160
	(0.019)***	(0.018)***	(0.026)***
Delinquencies	-0.018	-0.013	-0.022
	(0.005)***	(0.007)**	(0.009)**
Income_range($0 or undefined)			
$1 - $24,999	0.020	0.010	0.060
	(0.011)*	(0.014)*	(0.024)**
$25k - $49,999	0.029	0.022	0.065
	(0.009)***	(0.012)*	(0.023)***
$50k - $74,999	0.034	0.028	0.068
	(0.010)***	(0.012)**	(0.024)***
$75k - $99,999	0.024	0.010	0.081
	(0.011)**	(0.013)*	(0.026)***
$100k +	0.011	0.001	0.057
	(0.010)*	(0.013)*	(0.027)**
Employment _status (full time)			
Part-time	-0.040	-0.052	-0.040
	(0.014)***	(0.026)**	(0.027)**
Self employed	-0.010	-0.015	-0.012
	(0.023)	(0.019)*	(0.012)*
Information attributes			
Include_picture	0.024	0.019	0.037
	(0.005)***	(0.007)***	(0.010)***
Elaboration		0.014	-0.030
		(0.023)	(0.037)
Bid_count	0.001	0.001	0.001
	(0.000)***	(0.000)***	(0.000)***
Loan attributes			
Requested amount ($1000)	-0.025	-0.024	-0.029
	(0.002)***	(0.002)***	(0.003)***
SQ Requested amount	0.000	0.000	0.000
	(0.000)***	(0.000)***	(0.000)***
Offer_ interest_ rate	0.001	0.002	0.001
	(0.000)***	(0.000)***	(0.001)
N	10,281	7,249	2,983

Table 4-4: Predicted probabilities for different credit risk over a range of loan size

In this table we present predicted probabilities for different credit risk over a range of loan sizes conducted after the Probit regression.

Requested amount ($1,000)	AA	A	B	C	D	E	HR
1	69.7%	62.5%	58.6%	51.1%	43.9%	32.5%	22.7%
2	62.9%	55.4%	51.6%	44.2%	37.4%	27.1%	18.5%
3	55.8%	48.3%	44.6%	37.7%	31.5%	22.3%	15.0%
4	48.7%	41.6%	38.1%	31.7%	26.1%	18.2%	12.0%
5	41.9%	35.2%	32.1%	26.3%	21.5%	14.7%	9.7%
6	35.6%	29.5%	26.7%	21.7%	17.5%	11.8%	7.7%
7	29.8%	24.4%	21.9%	17.6%	14.1%	9.5%	6.2%
8	24.7%	20.0%	17.9%	14.3%	11.4%	7.6%	5.0%
9	20.2%	16.2%	14.5%	11.5%	9.1%	6.1%	4.1%
10	16.4%	13.1%	11.6%	9.2%	7.3%	4.9%	3.4%
11	13.2%	10.5%	9.3%	7.4%	5.9%	4.0%	2.8%
12	10.6%	8.4%	7.5%	5.9%	4.8%	3.3%	2.3%
13	8.5%	6.7%	6.0%	4.8%	3.9%	2.8%	2.0%
14	6.8%	5.4%	4.9%	3.9%	3.2%	2.3%	1.7%
15	5.5%	4.4%	4.0%	3.2%	2.7%	2.0%	1.4%
16	4.5%	3.6%	3.3%	2.7%	2.2%	1.7%	1.2%
17	3.7%	3.0%	2.7%	2.3%	1.9%	1.4%	1.1%
18	3.0%	2.5%	2.3%	1.9%	1.6%	1.2%	0.9%
19	2.5%	2.1%	1.9%	1.6%	1.4%	1.1%	0.8%
20	2.1%	1.8%	1.6%	1.4%	1.2%	0.9%	0.7%
21	1.8%	1.5%	1.4%	1.2%	1.0%	0.8%	0.6%
22	1.5%	1.3%	1.2%	1.0%	0.9%	0.7%	0.5%
23	1.3%	1.1%	1.0%	0.9%	0.8%	0.6%	0.4%
24	1.1%	1.0%	0.9%	0.8%	0.6%	0.5%	0.3%
25	1.0%	0.8%	0.8%	0.7%	0.5%	0.4%	0.3%

*All other variables are set at the mean

Table 4-5: Predicted probabilities for credit allocation at representative values

In this table we present predicted probabilities for credit allocation at represented values conducted after the Probit regression; with actual cases extracted from the population of loan requests.

Actual Cases	Probability of credit approval
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, homeowner, existing firm, image, elaborate	20%
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, rent home, existing firm, no image , elaboration	15%
Self-employed, Prime credit grade **AA**, income range $25k - $49,999k, no past due loans and no judgments, rent, new firm, no image, elaboration	12%
Full-time employed, Prime credit grade A, income range $25k - $49,999k, no past due loans and no judgments, homeowner, existing firm, image , elaboration	16%
Full-time employed, Prime credit grade **A**, income range $25k - $49,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	11%
Self-employed, Prime credit grade **A**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	10%
Full-time employed, Prime credit grade **B**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	9%
Self-employed **,** credit grade **C**, income range $25k - $49,999k, past due loans and judgments, home owner, existing firm, image, elaboration	7%
Full-time employed average credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	6%
Self-employed credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	6%
Full-time employed credit grade **E**, income range $25k - $49,999k, delinquencies, no judgment, homeowner, existing firm, image, elaboration	5%
Self-employed, credit grade **E**, income range $25k - $49,999k, delinquencies, judgment, rent, existing firm, image, elaboration	5%
Full-time employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, homeowner, new firm, image, elaboration	3%
Self-employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, rent home, new firm, no image, elaboration	2%
Self-employed, high risk credit grade **HR**, income range $1 - $24,999, past due loans and judgments, rent home, new firm, no image, no elaboration	1%

Requested amount and interest rates set at their mean values, $10,430 and 18.5 percent respectively

Table 4-6: Robustness Test Probability of funding – without credit grades

This table reports the Probit regression results for factors driving credit approval. The first two regressions present estimates for the general and parsimonious specifications for all loan requests. The last two regressions present estimates for parsimonious specifications for loans from existing firms only and for loans from new firms only. In all regressions, the dependent variable is binary taking the form 1 if credit the loan request was funded and 0 otherwise. The explanatory variables include owner, firm, and information attributes: home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan.

The controls are requested loan size, offer interest rate, employment status, and income. Time and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Pr\ (Approval|1) = \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively.

Variable	(1) General model (all)	(2) Parsimonious Model (all)	(3) Existing firms (only)	(4) New firms (only)
Constant	-0.429	-0.469	-0.370	0.355
	(-0.478)	(-0.531)	(-0.747)	(0.433)
Owner Attributes				
Home_owner	0.225***	0.221***	0.181***	0.373***
	(4.735)	(4.653)	(3.259)	(3.756)
Delinquencies	-0.300***	-0.288***	-0.257***	-0.323***
	(-5.967)	(-5.782)	(-4.402)	(-3.149)
Judgements	-0.347**	-0.342**	-0.337**	-0.430**
Income_range ($ 0 – unable to verify)				
$1 - $24,999	0.133	0.145	0.047	0.470**
	(1.267)	(1.387)	(0.376)	(2.048)
$25,000 - $49,999	0.194**	0.198**	0.126	0.472**
	(2.105)	(2.158)	(1.222)	(2.128)
$50,000 - $74,999	0.243***	0.241**	0.162	0.560**
	(2.584)	(2.575)	(1.549)	(2.434)
$75,000 - $99,999	0.137	0.138	-0.017	0.704***
	(1.282)	(1.291)	(-0.142)	(2.760)
$100 000 plus	0.033	0.039	-0.058	0.451
	(0.305)	(0.364)	(-0.499)	(1.638)
Employment_status (full time)				
Part-time	-0.446**	-0.439**	-0.462*	-0.461*
	(-2.512)	(-2.478)	(-1.943)	(-1.560)
Self-employed	-0.097*	-0.093*	-0.130**	0.113*
	(-1.843)	(-1.772)	(-2.229)	(0.886)
Firm Attributes				
Existing_firm	0.068			
	(1.343)			
Information Attributes				
Include_picture	0.220***	0.218***	0.162**	0.410***
	(4.108)	(4.085)	(2.541)	(3.767)

Table 4-6 continue

Variable	(1) General model (all)	(2) Parsimonious Model (all)	(3) Existing firms (only)	(4) New firms (only)
Elaboration	0.199	0.212	0.627*	-0.406
	(0.805)	(0.859)	(1.704)	(-0.910)
Bid_count	0.008***	0.008***	0.008***	0.012***
	(35.147)	(36.166)	(30.743)	(18.356)
Loan Attributes				
Requested_amount($1000)	-0.183***	-0.156***	-0.147***	-0.210***
	(-13.039)	(-24.677)	(-20.381)	(-13.881)
Offer_interest_rate (%)	0.004**	0.012***	0.012***	0.014***
	(-0.286)	(-4.861)	(-3.850)	(-2.695)
Time fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Observations	10,278	10,278	7,246	2,983
Pseudo R^2	0.449	0.449	0.439	0.518
Loglikelihood	-2028	-2030	-1488	-491.3

Table 4-7: Robustness test - marginal effects for factors driving probability of funding

This table presents the marginal effects after Probit models based on all variables set at their means. The marginal effects for categorical variables show how Pr (Approval = 1) is predicted to change as a particular factor variable changes from 0 to 1, holding all other independent variables at zero. The marginal effect of a continuous variable measures the instantaneous rate of change, which may or may not be close to the effect on Pr (Approval=1) of a one unit increase in the independent variable. *, **, *** stand for 0.1, 0.05, and 0.01 significance levels respectively.

	All loans	Existing firms (only)	New firms(only)
Owner attributes			
Home_owner	0.024	0.021	0.034
	(0.005)***	(0.006)***	(0.009)***
Delinquencies	-0.031	-0.029	-0.029
	(0.005)***	(0.007)***	(0.009)***
Judgements	-0.037	-0.038	-0.039
	(0.015)**	(0.017)**	(0.030)**
Income_range($0 or undefined)			
$1 - $24,999	0.015	0.005	0.043
	(0.011)*	(0.014)*	(0.021)**
$25k - $49,999	0.021	0.014	0.043
	(0.009)**	(0.012)**	(0.020)**
$50k - $74,999	0.026	0.018	0.051
	(0.010)***	(0.012)*	(0.021)**
$75k - $99,999	0.014	-0.002	0.064
	(0.011)**	(0.014)*	(0.023)***
$100k +	0.004	-0.007	0.041
	(0.010)	(0.013)	(0.025)
Employment Status (full time)			
Part-time	-0.041	-0.052	-0.042
	(0.014)***	(0.027)*	(0.027)*
Self-employed	-0.010	-0.015	0.010
	(0.006)*	(0.007)**	(0.012)*
Information attributes			
Include_picture	0.023	0.018	0.037
	(0.006)***	(0.007)**	(0.010)***
Bid_count	0.001	0.001	0.001
	(0.000)***	(0.000)***	(0.000)***
Loan attributes			
Requested amount ($1000)	-0.017	-0.017	-0.019
	(0.001)***	(0.001)***	(0.001)***
SQ Requested amount	0.000	0.000	0.000
	(0.000)***	(0.000)***	(0.000)***
Offer_ interest_rate	0.001	0.001	0.001
	(0.000)***	(0.000)***	(0.000)***
N	10,278	7,246	2,983

Table 4-8: Robustness test Tobit estimation of credit approval decision

This table report regression results for factors driving credit approval. In all regressions, the dependent variable is a percentage of the loan funded denoted by ($\frac{\text{funded loan amount}}{\text{request loan amount}}$). The explanatory variables include credit grade[+], home owner, firm, and information attributes: home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan. The controls are requested loan size, offer interest rate, employment status, and income. Time and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Fraction\_funded = \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively

	Fraction Funded With credit grade	Fraction Funded Without credit grade
Constant	0.499***	0.379**
	(2.871)	(2.073)
Owner Attributes		
Home_owner	0.009	0.032***
	(1.265)	(4.444)
Credit_grade (ref AA)		
A	-0.060***	
	(-4.700)	
B	-0.134***	
	(-10.224)	
C	-0.231***	
	(-16.781)	
D	-0.307***	
	(-19.614)	
E	-0.390***	
	(-21.514)	
HR	-0.511***	
	(-30.005)	
Delinquencies	-0.047***	-0.104***
	(-6.436)	(-14.020)
Judgements	-0.011	-0.054***
	(-0.650)	(-3.102)
Income_range ($ 0 – unable to verify)		
$1 - $24,999	0.061***	0.044***
	(4.172)	(2.870)
$25,000 - $49,999	0.088***	0.068***
	(7.163)	(5.342)
$50,000 - $74,999	0.108***	0.092***
	(8.599)	(7.094)
$75,000 - $99,999	0.109***	0.090***
	(7.640)	(6.018)
$100 000 plus	0.089***	0.079***
	(6.596)	(5.654)
Employment_status (full time)		
Part-time	-0.071***	-0.056**
	(-2.982)	(-2.273)
Self-employed	-0.050***	-0.047***
	(-6.936)	(-6.204)
Firm Attributes		
Existing_firm	0.016	0.025
	(2.198)	(3.272)

Table 4-8 continue

	Fraction Funded With credit grade	Fraction Funded Without credit grade
Information Attributes		
Include_picture	0.033***	0.039***
	(4.662)	(5.159)
Elaboration	-0.026	-0.007
	(-0.811)	(-0.223)
Bid_count	0.002***	0.003***
	(74.880)	(85.116)
Loan Attributes		
Requested_amount($1000)	-0.020***	-0.014***
	(-38.560)	(-29.219)
Offer_interest_rate (%)	0.005***	0.001**
	(12.664)	(1.285)
Time fixed effects	YES	YES
Industry fixed effects	YES	YES
Number of observations	10,303	10,303
Pseudo R^2	0.569	0.492
Loglikelihood	-3315	-3911

Figure 4-1: Geographic distribution of loan requests

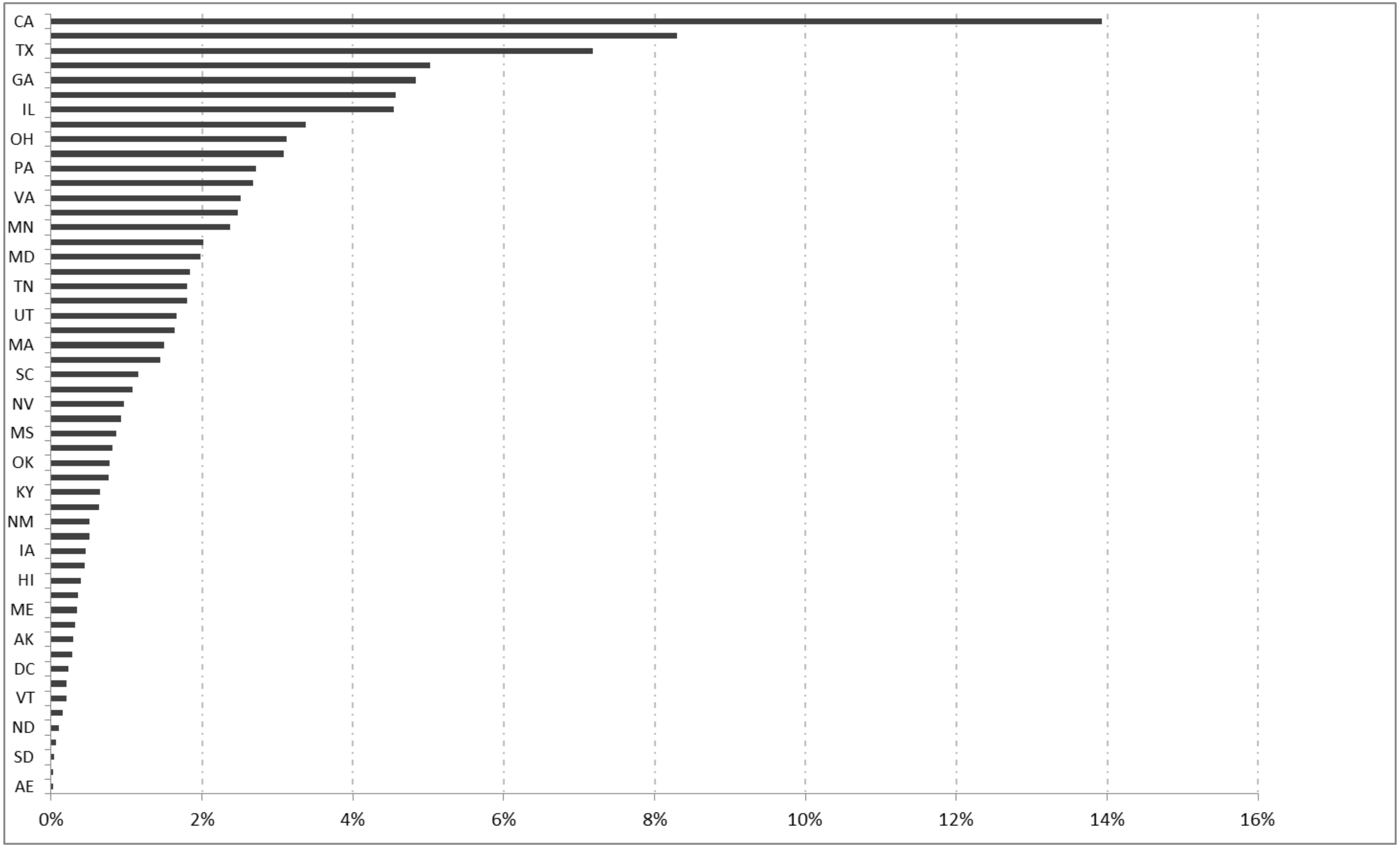
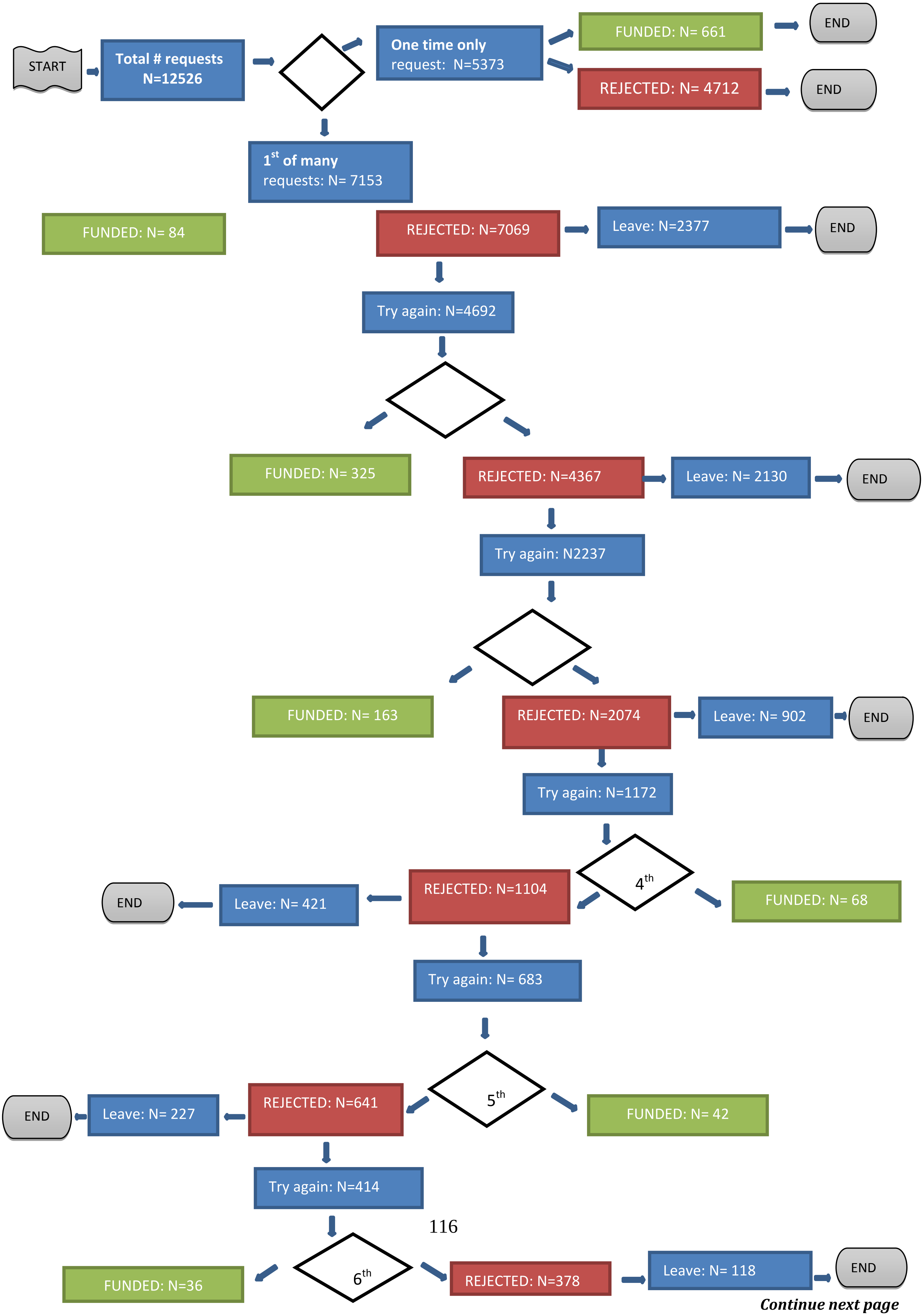


Figure 4-2: Flow chart showing the attempts of getting funding



Continue next page

Continue from previous page

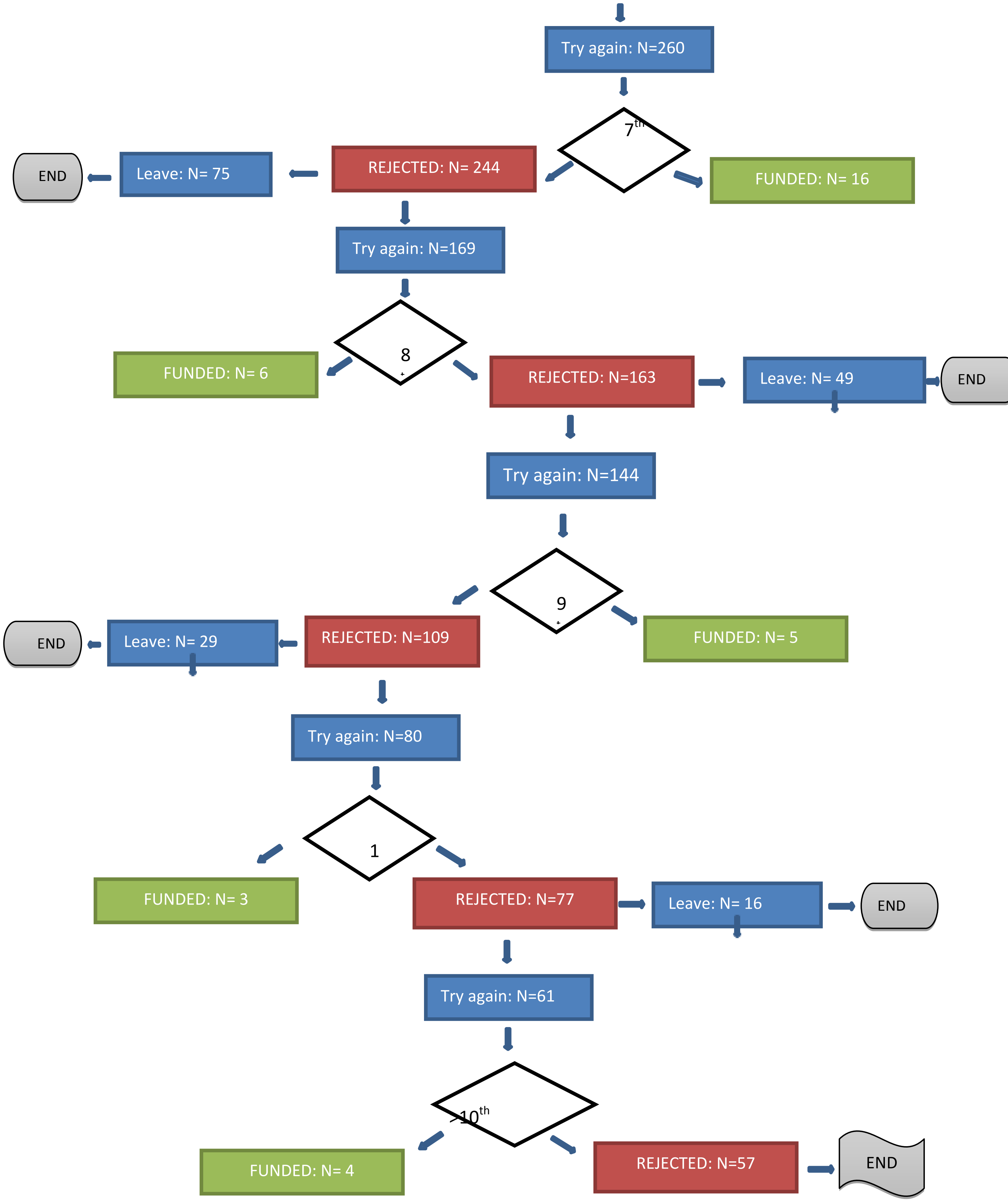


Figure 4-2: Square measurement

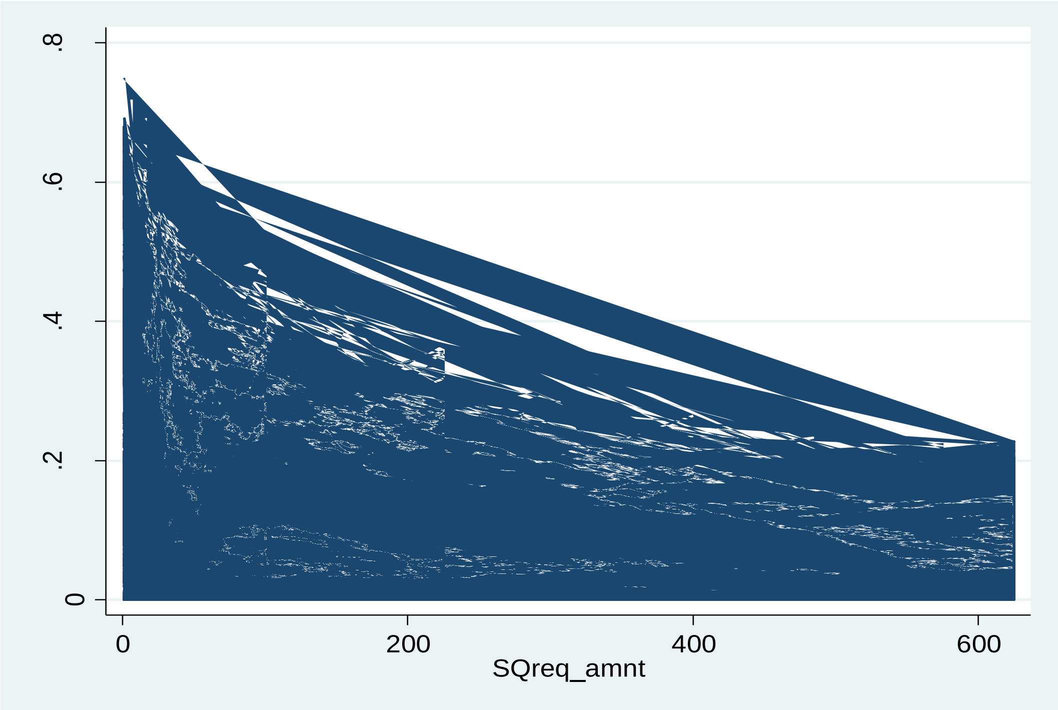


Figure 4-3: Predictive values across different loan size for borrowers in different credit risk categories

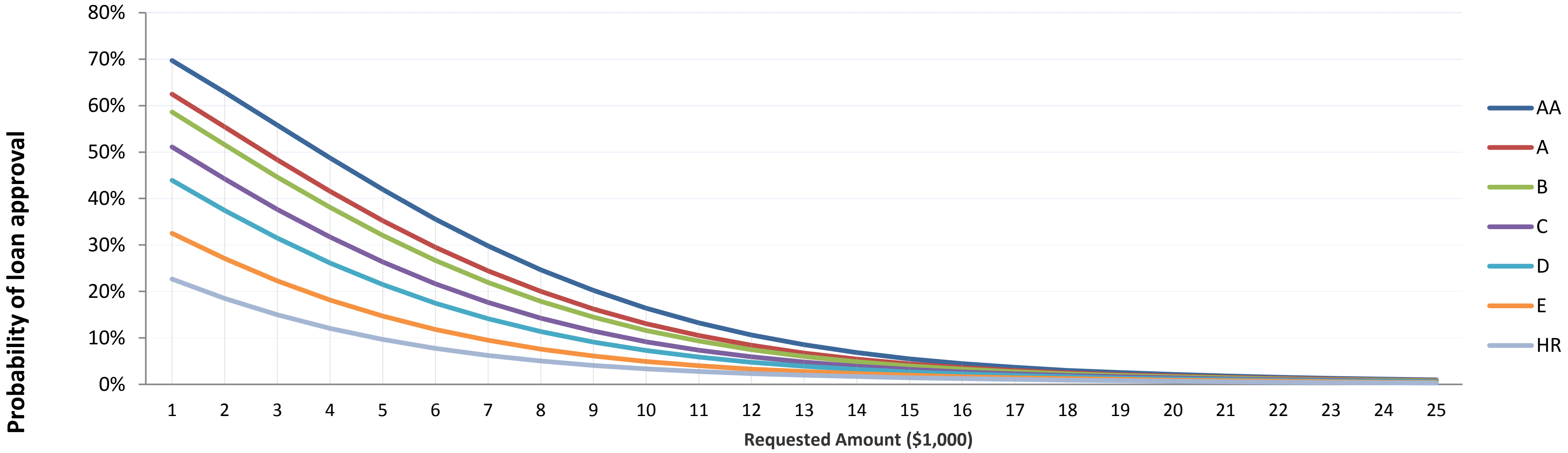
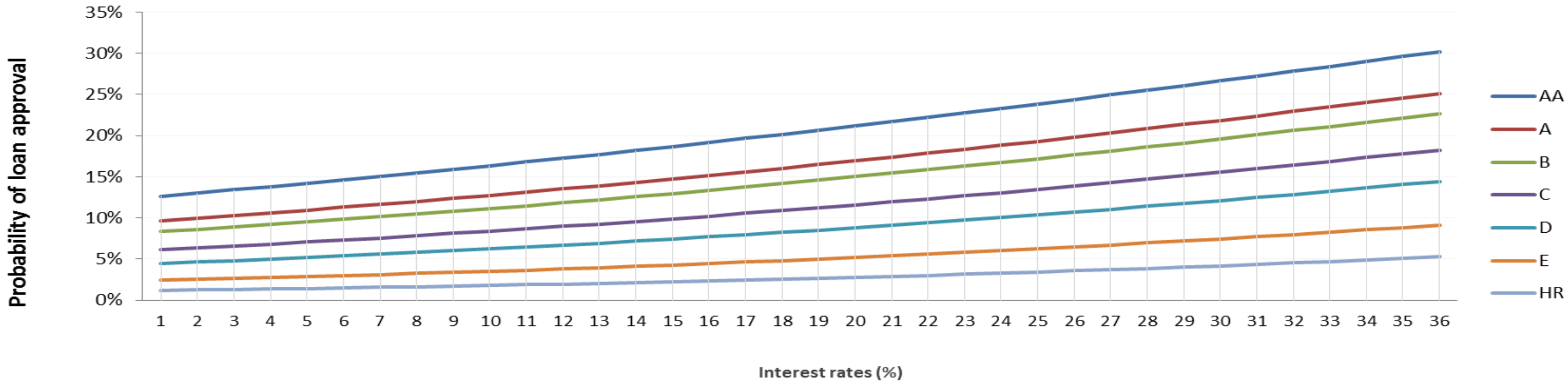


Figure 4-4: Predictive values across different loan size for borrowers in different credit risk categories



Chapter 5 Factors driving the cost of credit in P2P lending

1.0 Introduction

The first empirical study, in the previous Chapter, examined the determinant effects of observable borrower and firm characteristics and information variables on the likelihood of credit allocation. As stated earlier, credit allocation is not only about access to credit but also about cost of credit. Hence, it is the purpose of this second study to examine empirically the fundamental determinants of interest rates for small business loans in the P2P lending context. The chapter focuses on whether the variations in interest rates paid by small business borrowers can be explained by differences in the observable borrower and firm characteristics and information variables (Diamond, 1989; Keasey and Watson, 2000; Petersen and Rajan, 1994; Berger and Udell, 1995; Harhoff and Körting, 1998; Cowling, 1999; Hanley and Crook, 2005; Burke and Hanley, 2006).

Our empirical tests are based on the premise that: if the methods we identified from theory, typically adopted in traditional lending, such as home ownership, credit grades, credit history, reputation built over time, and firm age help attenuate information problems in the P2P lending context, then we should observe higher risk small business borrowers (those less likely to repay the credit) associated with higher interest rates and lower risk small business borrowers (those likely to repay credit) associated with lower interest rates. This evidence would suggest that: either little risk is generated by small business borrowers with better observable borrower and firm characteristics; or that the risk that is generated is generally offset by lenders' ability to select risk.

In addition, to the extent that the mechanisms specific to P2P lending such as the optional pictures and text elaboration may possibly generate information about small business borrowers (or the business ventures) and perhaps attenuate information problems, we also explore whether inclusion of this kind of information result in cheaper (expensive) credit for small business borrowers.

Our empirical analysis also distinguishes new business start-ups from established firms. New business start-ups are arguably more risky. Hence, we test the following hypotheses:

H1b small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are more likely to pay lower interest rates

H2b small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are more likely to pay lower interest rates

H3b small business borrowers who have successfully paid back a previous loan are more likely to pay lower interest rates

H4b existing firms are more likely to pay lower interest rates (relative to new business start-ups) in the P2Plending context

H6b small business borrowers who use text elaborations are more (less) likely to be pay cheaper interest rates in the P2P lending context

H7b small business borrowers who have previous failures are more likely to pay higher interest rates

H8b small business owners who post pictures are more (less) likely to pay lower interest rates in the P2P lending context

The rest of the Chapter is organized in the following manner. We begin by providing summary statistics and simple bivariate correlations of variables. Then in section 5.2, estimates based on multivariate analysis of the determinants driving the price of small business credit on Prosper are presented. Finally, in section 5-3, we present our robustness checks.

5.1 Descriptive analysis

Table 5-1 presents descriptive statistics relating to the 1417 observations that resulted in funded loans for the period 2007 - 2013. The table indicates that across all funded loans, the mean interest rate offered and the final interest rate paid by small business borrowers were 21.3 and 18.5 percent respectively with a considerable standard deviation. Figure 5-1 displays how the offer interest rate and final interest rates vary across time. We observe three things from the figure: first the interest rates offered seem to have a selection effect - lenders appear cautious of borrowers offering higher interest rates; opting to ration credit instead. Second, we observe that over time, the final interest rate paid approaches the (higher) levels of the offer interest rates. Third, compared to typical small business loans from traditional lending institutions like banks, loans on Prosper, on average, are more expensive. According to SBA data (2010), typical small business loans (under $100,000) from US banks show a national average of 13.2 percent;

unsecured loans from Credit unions averaged 12.8 percent per annum for the period 2006 - 2009, whilst business credit cards had a national average of 15 percent.

...........Table 5-1 goes around here...........

Interestingly, when we break down the interest rates paid by credit grade as shown in Table 5-1, we observe that for small business borrowers in the low risk categories (AA) Prosper offers even better interest rates in comparison to business credit cards (11.7 percent vs. 15 percent); fixed term unsecured bank loans (11.7 percent vs. 13.2 percent) and fixed term unsecured Credit union loans (11.7 percent 12.8 percent). This observation may seem to suggest that for a subset of low risk small business borrowers, this market may be competing with traditional banks. All other borrowers in the remaining credit grade categories seem to pay interest rates higher than those available in traditional debt lending; more comparable to returns earned by equity finance investors as shown in Table 5-2 (Damodaran, 2010). In fact we observe from Table 5-2 that that interest that high risk borrowers in credit grades HR and E seem somewhat comparable to returns from Venture Capitalists in 2007, which may seem to suggest that for a subset of small business borrowers P2P lending may actually be an expensive form of finance.

...........Table 5-2 goes around here...........

5.2 Regression analysis

Next in order to confirm the suggested associations as suggested by the univariate results, we conducted multivariate analysis. The dependent variable used in the analysis is the interest rate paid by the small business borrowers (Interest rate). Following previous studies (Keasey and Watson, 1; Cowling, 1999; Cressy and Toivanen, 2001; Hanley and Crook, 2005; Burke and Hanley, 2006) the interest rate was adjusted according to the prime rate to provide the interest rate premium so as to allow for a reliable comparison between data collected at different points in time. When lenders on Prosper grant a loan to small business borrowers they gather information to assess the credit risk of their projects. We argue that rational lenders in the P2P context will accept an interest rate equal to the underlying cost of funds plus a premium that will depend on the quality of the credit risk presented by the borrower. As indicated in the methodology section (Chapter 3), in the regression analysis, we control for observable borrower and firm characteristics that proxy for firm quality as well as the prevailing conditions in the market when the loan was made. We include variables reflecting the main characteristics of the small business borrower: home ownership, credit grade, credit history variables - delinquencies and judgements, employment status and income. We include variables reflecting firm attributes namely: age of the firm (denoted as established or new firm) and dummies for firm industry.

We also include information variables: given that lenders on Prosper can see whether a small business borrower has any previous loans, we include dummy (repeat loan), signifying whether the finance requested is follow-up finance or whether the loan represents the first ever granted request. We also add information variables from the (optional) pictures and text elaboration. Finally, we add some controls for the main features of the loan: loan size, offer interest rate; dummies for the region where the small business borrower is located; and time dummies to control for variation in the market conditions. The regressions estimated by Tobit follow the form:

$$\begin{aligned} \textit{Interest rate} = \alpha + \beta_1\, Owner_i + \beta_2\, Firm_i + \beta_3\, InformationAttributes_i + \beta_4\, LoanAttributes_i \\ + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\, Region_i + \mu \qquad (5\text{-}1) \end{aligned}$$

The empirical results are presented in Table 5-3; which is made up of 3 columns. In the first column, we report results for the general regression specification; including all the regressor variables that could possibly drive the cost of credit on Prosper (identified from literature). We tested down all regressor variables to derive a parsimonious model shown in column 2; dropping variables from the general estimation which were insignificant at the 10 percent level ($p > 0.10$). The coefficients of the included variables may be interpreted as reflecting the association between the included variables and the risk of the loan as reflected in its price. The results of the Likelihood Ration (LR) test confirm that the parsimonious model is a better estimation; the null hypothesis that the excluded regressors collectively have no role in predicting our dependent variable is decisively accepted (LRX^2= - 3562, *df* = 102, *p* <0.10). The third columns subsequently report the determinants driving the cost of credit separate for when we exclude the credit variable; on the basis that it may be capturing some of the effects we observe from delinquencies and home ownership.

...........Table 5-3 goes around here...........

On the whole, the general findings from this study suggest that at an average lending of between 18 percent and 20 percent; P2P lending is a very expensive form of debt finance. Banks typically refuse to extend credit given such high interest rates as this tends to alter the borrower pool as put forward by theory (see Stiglitz and Weiss, 1981; Bosanko and Thakor,1987) such that only the riskiest of borrowers have projects that generate returns that are high enough to be able to re-pay these interest rates. Consequently, if we were to characterise P2P lending we would effectively conclude that it is typically a high cost finance with required returns expected to be likely in the levels of Business Angels and VC equity investments (Bygrave and Quill,2007).

Further insight we get from our second study is that borrower reputation, stipulated by credit grades is the single most important determinant of the cost of credit .The significance of using the credit grade helps to reduce the problem of moral hazard. The cost of defaulting will result in poorer scores – quantifiable to an increase of 80 basis points. Second, we see from our results that collateral, which was such an important determinant in reducing moral hazard in Stiglitz-Weiss theory, in the P2P lending it is unimportant. The fact that P2P lending relaxes collateral requirements and the fact that reputation is the single most important variant in assessing the cost of credit on P2P websites, highlights some of the facts that were not previously considered by previous theory. That is, moral hazard may be solved by other means in the form of credit scores.

More specifically, we find that small business borrowers in high risk credit categories (HR) will pay as high as 500 percent more in interest rate when compared to those in premium credit grade categories (AA). This suggests that 'good types' defined in terms of credit rating, get their loans at a lower interest rate (H_{2b}). These results may be driven by the fact that either little risk is generated by small business borrowers with better creditworthiness, or that the risk that is generated by higher risk borrowers is generally offset by the lenders charging higher interest rates for it. Given that in the previous empirical analysis (Chapter 4) we identified that the credit grade variable may also include information already captured by the *Home_ownership* variable and other credit history variables, in columns 3 of Table 5-3 we estimate the Tobit regressions based on the estimation defined in equation 5-1, excluding the credit grade variable. Indeed we observe that the *Home_ownership* variables become significant such that small business borrowers who own homes will pay 3 percent less in interest rate per $1,000 extended by lenders relative to those who rent homes (H_{1b}). Similarly, borrowers who have failed to honour previous loan commitments end up with higher interest rates (H_{7b}).

In terms of firm attributes, we observe from column 2 of Table 5-3 that the coefficient on our variable of interest *Existing_ firm*, is not significant (H_{4b}). Thus, we do not find evidence that lenders on Prosper incorporate firm level characteristics when pricing loans. This result seem to suggests that the pricing of loans in this context may possibly be relatively idiosyncratic - the interest rate on the funded loans may depend more on personal reputation of the small business owner than on the observed characteristics of the firm. This finding seems counter to results found in small finance literature (see Keasey and Watson; 2000; Cressy and Toivanen, 2001; Cowling, 1999) where firm age is one of the key determinants of pricing credit. Additionally, we also find that our industry variables are not significant (not reported).

Interestingly, in terms of the information variables: the insignificant coefficient of the variable *Repeat_loans* suggests that building a relationships in this context may not necessarily translate to cheaper credit as suggested by model of Petersen and Rajan, 1994; hence we do not find support for $H_{3b.}$ This is hardly surprising given that relationships are weak at best.

However, the negative and statistically significant coefficient on the *Include_picture* variable suggests that by including a picture in the loan requests, small business borrowers will pay almost 50 percent less in their final interest rates per $1,000 of their funded loan relative to those who do not include a picture (H_{8b}). This result appears to suggest that lenders seem to value pictures as a mechanism of reducing information asymmetries - even though these pictures are not verified by Prosper.

In terms of control variables, we find that small business lenders, who offer high interest rates, will end up paying more for credit on Prosper. Interestingly, the loan size variable *Requested_ amount* has a positive and significant sign, suggesting that P2P lenders expect to be compensated for the risk of extending larger loan amounts; hence in this context larger loans are expensive. This observation seem counter intuitive to theory and what typically happens in traditional lending, where larger loans are typically associated with lower interest rates (Besanko and Thakor, 1987; Cressy, and Toivanen,2001; Hanley and Crook, 2006); given the fact that they have more collateral. But given the fact that P2P loans are not secured by collateral, high interest rates may induce moral hazard behaviour (Stiglitz and Weiss, 1981).

In Table 5-4 we move on to present estimates for Tobit regressions separately for established firms and for new business start-ups in order to facilitate comparison as to whether factors driving interest rates differ between these two groups.

$$Interest\ rate\_Existing\_Firm = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ InformationAttributes_i + \beta_4\ LoanAttributes_i + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\ Region_i + \mu \tag{5-2}$$

$$Interest\ rate\_New\_Business = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ InformationAttributes_i + \beta_4\ LoanAttributes_i + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\ Region_i + \mu \tag{5-3}$$

The findings for factors driving cost of credit for established firms substantially mirror those of new business start-ups and those found for all firms; further validating the earlier observed result, evident from the *Existing_firms* variable, suggesting that lender on Prosper may not necessarily pay attention to firm characteristics when deciding on the cost of credit.

...........Table 5-4 goes around here...........

We now move on and try and illustrate the impact of the statistically significant factors that were shown above to influence interest rates paid. In Table 5-5 we compile predicted probabilities for factors driving interest rate at representative values. The baseline case is an individual with Prime credit grade A, income range \$25k - \$49,999k, with previous delinquencies and judgements, who rents their home, running an existing business, who opts not to include a picture but gives a text elaboration - this borrower has an 11 percent likelihood of funding success and is likely to pay approximately 13.7 percent interest rate. Compared to the baseline case, a borrower with a premium credit grade AA, full time employed income range $50k - $74,999k, no past due loans, who is a homeowner, who includes a picture has a 20 percent chance of getting a loan request funded at a cost of almost 12.5 percent interest rate; this translates to a probability of funding that is twice as likely than the base case, with an interest rate that is only 1.2 percentage point less. This renders the credit grade as an influential determinant of the price. Whilst, compared to the baseline case, a high risk borrower with a HR credit grade, full time employed, income range $75k - $99,999k with delinquencies and who rents their home, has a 1 percent chance of being funded and can expect to pay as high as 30 percent interest rate. Being delinquent in the past on loan obligations only reduce the probability of funding by 0.56 percentage points which seems to support our argument that lenders in this market are forgiving.

...........Table 5-5 goes around here...........

Our predictions do confirm that all in all including a picture will increase funding success and lower interest rates. When we compare otherwise identical cases in terms of credit grades, differing only in inclusion of pictures (as shown in the first two rows of Table 5.5), we observe that the borrower who includes a picture can save 0.5 percentage points in interest rates. Unfortunately, for the highest risk category (HR) even inclusion of a picture simply would not be enough.

All in all, our results suggest that prospective borrowers with low credit scores, it will be cheaper to partner with a small business owner who has a good credit grade in order to gain access to the market.

5.3 Robustness Check

In the preceding sections, we found that credit grades, previous credit history and the inclusion of pictures significantly impacts the cost of credit on Prosper; whilst the remainder of our explanatory variables namely: repeat loan, firm variables and stories that small business borrowers tell were not important. In this subsection we perform several additional regressions as robustness checks.

5.3.1 Changes in the method of estimation – accounting for selection bias

Our previous estimates on factors driving interest rates were based on a sample of 1417 funded loans only; a sample that is not randomly selected. Because we do not estimate factors driving cost of credit for the population (i.e. randomly selected sample) our results may be biased. This challenge emphasises the need to model sample selection explicitly. Hence as a robustness check, we have adopted a 2-stage Heckman model to estimate the interest rate equation accounting for selection bias as follows:

Stage 2 Outcome regression: observed only if loan funded (i.e. Pr *(Approval)* = 1)

$$\begin{aligned}\textit{Interest rate} &= \alpha + \beta_1\, Owner_i + \beta_2\, Firm_i + \beta_3\, InformationAttributes_i + \beta_4\, LoanAttributes_i \\ &\quad + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\, Region_i + \mu \qquad (5.3a)\end{aligned}$$

Stage 1 selection regression: probability of credit approval is 1 if loan funded, and 0 otherwise

$$\begin{aligned}Pr\ (Approval) &= \alpha + \beta_1 Owner_i + \beta_2 Firm_i + \beta_3 InformationAttributes_i + \beta_4 LoanAttributes_i \\ &\quad + \beta_5 Industry_i + \beta_6 Macro_i + \beta_7\, Region_i + \varepsilon \qquad (5.3b)\end{aligned}$$

Table 5-6 report our result. The estimated coefficients in this case are an indicator of determinants driving interest rates net the observed selection bias. All in all, we observe the same pattern of results, with the exception of home ownership, meaning that after controlling for selection bias, it seems credit grades are still important determinants of interest rates. Similar to the Tobit, we find that borrowers with delinquencies also pay higher interest rates ($p<0.01$); whilst small business borrowers that include a picture ($p<0.01$). We find however that the Home ownership variable now is positive and statistically significant; suggesting that borrowers who own homes may appear as better risk to lenders. These results seem to suggest that even after selection; lenders in this context may still be unable to resolve information asymmetries such they seem to compensate by charging higher interest rates for those in high risk credit grades (C, D E and HR) and look to home ownership, previous credit history and inclusion of pictures as signals which may help attenuate information issues.

...........Table 5-6 goes around here...........

In terms of firm attributes, we observe that our variable *Existing_firm* is still insignificant. Moreover, consistent with our previous observation, repeat loans do not necessarily translate to cheaper credit in the P2P lending context. We also find that the stories that prospective borrowers tell do not seem to affect price of credit.

In terms of our loan attributes, we observe from Table 5-6 that it seems the result observed from the Tobit estimation on the *Requested_amount* variable and *Offer_interest_rate* still hold after accounting for selection bias. All else equal, small business borrowers who ask for larger loans, who offer higher interest rates do end up paying high prices for credit in the P2P lending context. Moreover, our control variables *Employment_status,* and *Income_range* variables included are significant in the selection equation, but not in the outcome regression; suggesting that once the lenders have looked at the prospective borrower's income and employment status when making a decision of whether or not to extend credit - these variables don't seem to influence the interest rate borrowers pay.

Finally, in term of justifying whether adoption of the Heckman estimation is reasonable, we see from Table 5-6 that the LR test of independence is supported. Further, the Null hypothesis of zero correlation between the error terms of the selection (equation 5-4b) and outcome (equation 5-4b) regressions was rejected at the 0.05 level; suggesting that our results can indeed be generalised to the population (of funded and declined loans).

5.3.2 Including ex post default

Next we, we looked at the robustness of our results based on *ex post* default; so as to establish whether indeed the high risk borrowers were paying higher prices for loans. From our results in Table 5-7, the ability of P2P lenders to ascertain the small business borrower's real 'revealed' quality is evident in the positive and significant coefficient of our variable *ex_default* ($p<0.01$).

Here we observe that small business borrowers who were higher risk (unidentifiable upfront) were in fact charged higher interest rates, perhaps in anticipation of this deterioration of their credit quality. All in all, the estimated coefficients for our key explanatory variables are generally consistent with our previous findings.

...........Table 5-7 goes around here...........

5.4 Chapter Summary

Recalling the research question, this study has empirically examined the determinants of interest rate paid by small business borrowers on Prosper .Given that the primary goal for this study was not so much to get the best estimation of loan rate; rather, it was to find the estimation implicitly used by the ‘less sophisticated’ Prosper lenders in pricing credit. This type of an enquiry helps us determine if mechanisms typically used in traditional lending can help attenuate information issues in pricing credit within a new context.

All in all, there is evidence that the interest rate paid by small business borrowers in this context incorporate a number of easily observable specific risk/cost characteristics of the borrower. In addition, there is also evidence that interest rate reflect the reduction in information asymmetries due to inclusion of pictures and those who are home owners. We find however that firm specific information does not affect the pricing of loans. This result seem to suggests that the cost of loans in this context may possibly be relatively idiosyncratic - the interest rate on the funded loans may depend more on personal reputation of the small business owner than on the observed characteristics of the firm.

In sum, main insight we get from this study is that borrower reputation, stipulated by credit grades is the single most important determinant of the cost of credit .The significance of using the credit grade helps to reduce the problem of moral hazard. Second, we see from our results that collateral, which was such an important determinant in reducing moral hazard in Stiglitz-Weiss theory, in the P2P lending it is unimportant. The fact that P2P lending relaxes collateral requirements and the fact that reputation is the single most important variant in assessing the cost of credit on P2P websites, highlights some of the facts that were not previously considered by previous theory. That is, moral hazard may be solved by other means in the form of credit scores

Table 5-1: Summary statistics of cost of capital

This table contains the descriptive statistics of the cost of capital based on all funded loans. We report the mean and the standard deviation of each variable.

Funded Loans	Obs	Mean interest rate (%)	Standard Deviation	Min	Max
Requested loan ($1,000)	1417	7.9	6.3	1	25
Offer Interest rate	1417	21.3	9.0	4.5	36
Final Interest rate	1417	18.5	8.6	4.3	36
Credit Grade					
AA	314	11.3	4.6	4.9	36
A	261	14.7	6.8	4.3	36
B	253	18.2	6.4	5.0	36
C	204	21.5	7.1	10	36
D	213	24.6	5.6	9.0	36
E	105	31.0	3.1	17	36
HR	67	31.5	4.8	15	36
Home owner	818	17.9	8.1	4.9	36
Rent	599	20.5	9.0	4.3	36
Firm Age					
Firm Exist	1027	18.8	8.7	5.6	36
Firm New	390	19.6	8.5	4.3	36
Industry					
agriculture	22	13.4	6.1	7.5	36
construction	14	18.7	5.5	9.0	36
finance, insurance, real estate	279	16.3	8.5	4.3	36
manufacturing	64	18.8	8.0	6.0	36
retail trade	420	20.9	8.4	6.0	36
services	572	18.6	8.4	5.0	36
wholesale trade	18	18.2	9.2	8.0	36
transportation	28	19.1	8.0	9.0	36
Include Picture	790	18.1	8.1	4.3	36
No Picture	627	20.1	9.1	5.0	36
Elaboration	1369	19.0	8.6	4.3	36
No elaboration	48	18.0	9.5	6.0	36
Repeat loans					
Once only	661	18.2	8.5	5.0	36
Twice only	84	19.0	8.9	4.3	36
Three times and more	672	20.4	8.0	5.6	36

Table 5-2: Returns from Venture Capitalists in 2007:

Reported rates of return for venture capital investment

	Three Years	Five Years	Ten Years	Twenty Years
Early/Seed VC	4.9%	5.0%	32.9%	21.5%
Balanced VC	10.8%	11.9%	14.4%	14.7%
Later VC	12.4%	11.1%	8.5%	14.5%
All VC	8.5%	8.5%	16.6%	16.9%

Data Source: Damodaran (2010): The Dark Side of Valuation: Valuing Young, Distressed, and Complex Business

Table 5-3: Figure 5 3: Factors driving cost of credit on Prosper

This table reports the Tobit estimates for factors driving credit on Prosper.com. The first regressions (column 1) present estimates for the general specifications for all loan requests. Next, column (2) present estimates for parsimonious specifications for all loan requests while column (3) represents estimates excluding the credit grade variable. In all regressions, the dependent variable is the interest rate (less prime rate). The explanatory variables include owner, firm, and information attributes: credit grade (excluded in column 3), home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan.

The controls are requested loan size, offer interest rate, employment status, and income. Time, Region and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Interest\ rate = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i + u_i$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively. t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	(1) General All	(2) Parsimonious All	(3) Parsimonious All Excluding credit grade
Constant	4.756	4.664	6.628
	(1.197)	(1.181)	(1.621)
Owner attributes			
Homeowner	0.077		
	(0.392)		
Credit grade (AA)			
A	0.750**	0.455	
	(2.533)	(1.595)	
B	1.730***	1.139***	
	(5.012)	(3.617)	
C	3.131***	2.350***	
	(7.861)	(6.824)	
D	3.915***	2.980***	
	(8.485)	(7.692)	
E	6.118***	5.143***	
	(10.756)	(10.142)	
HR	6.309***	5.304***	
	(10.142)	(9.652)	
Delinquencies	0.131		0.081*
	(0.604)		(0.404)
Judgements	0.375		0.575*
	(1.586)		(2.442)
Employment_Status(full time)			
Part-time	-0.611		
	(-0.792)		
Retired	-0.112		
	(-0.185)		
Self-employed	0.364		
	(1.605)		
Income_Range ($ 0 – unable to verify)			
$1 - $24,999	-0.654	-0.683	Table 5-3 continues
	(-1.293)	(-1.383)	(-0.516)
$25,000 - $49,999	-1.197***	-1.110***	-0.764*
	(-2.748)	(-2.723)	(-1.829)
$50,000 - $74,999	-1.184***	-1.016**	-0.755*
	(-2.673)	(-2.461)	(-1.782)

Table 5-3 continues

	(1) General All	(2) Parsimonious All	(3) Parsimonious All Excluding credit grade
$75,000 - $99,999	-1.378***	-1.140***	-0.898**
	(-2.905)	(-2.583)	(-1.963)
$100 000 plus	-1.483***	-1.113**	-1.109**
	(-3.129)	(-2.537)	(-2.430)
Firm Attributes			
Existing firms	-0.064		
	(-0.326)		
Information Attributes			
Repeat loans	-0.404		-0.221
	(-2.117)		(-1.128)
Include_picture	-0.456**	-0.457**	-0.567**
	(-2.105)	(-2.095)	(-2.559)
elaboration	0.194		
	(0.316)		
Loan Attributes			
Requested amount ($1000)	0.021**	0.062*	0.016*
	(0.404)	(0.221)	(0.302)
SQRequested amount	0.003		
	(1.267)		
Offer interest rate (%)	0.539***	0.648***	0.747***
	(9.524)	(12.703)	(15.227)
Time fixed effects	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes
Regional effects	Yes	Yes	Yes
Number of Observations	1,390	1,391	1,390
Pseudo R^2	0.291	0.289	0.279
Log Likelihood	-3546	-3562	-3682

Table 5-4: Predicted margins at represented values

In this table we present predicted probabilities of interest rate at represented values conducted after 2- stage Heckman regressions; with actual cases extracted from the population of loan requests

Actual Cases		Interest rate
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, homeowner, existing firm, image, elaborate	.20	12.5%
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, rent home, existing firm, no image , elaboration	.15	13.0%
Self-employed, Prime credit grade **AA**, income range $25k - $49,999k, no past due loans and no judgments, rent, new firm, no image, elaboration	.12	13.3%
Full-time employed, Prime credit grade A, income range $25k - $49,999k, no past due loans and no judgments, homeowner, existing firm, image , elaboration	.16	13.5%
Full-time employed, Prime credit grade **A**, income range $25k - $49,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	.11	13.7%
Self-employed, Prime credit grade **A**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	.10	13.7%
Full-time employed, Prime credit grade **B**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	.09	14.5%
Self-employed **,** credit grade **C**, income range $25k - $49,999k, past due loans and judgments, home owner, existing firm, image, elaboration	.07	15.6%
Full-time employed average credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	.06	16.5%
Self-employed credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	.06	16.9%
Full-time employed credit grade **E**, income range $25k - $49,999k, delinquencies, no judgment, homeowner, existing firm, image, elaboration	.06	19.7%
Self-employed, credit grade **E**, income range $25k - $49,999k, delinquencies, judgment, rent, existing firm, image, elaboration	.05	20.1%
Full-time employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, homeowner, new firm, image, elaboration	.03	22.9%
Self-employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, rent home, new firm, no image, elaboration	.02	27.3%
Self-employed, high risk credit grade **HR**, income range $1 - $24,999, past due loans and judgments, rent home, new firm, no image, no elaboration	.01	30.1%

Note: Requested loan amount set at the mean $ 10,430

Table 5-5: Factors driving cost of credit on Prosper – split by Firm Status

This table reports the Tobit estimates for factors driving credit on Prosper.com. The first regressions (column 1) present estimates for the general specifications for all loan requests. Next, column (2) and column (3) regressions present estimates for parsimonious specifications for loans from existing firms only and for loans from new firms only respectively. In all regressions, the dependent variable is the interest rate (less prime rate). The explanatory variables include owner, firm, and information attributes: credit grade, home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan.

The controls are requested loan size, offer interest rate, employment status, and income. Time, Region and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Interest\ rate = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i + u_i$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively. t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	(1) General All	(2) Existing firms Only	(3) New firms only	(4) Existing firms only (no credit grade)	(5) New firms Only (no credit grade)
Constant	4.756	1.084	4.147	2.849	6.287
	(1.197)	(0.271)	(1.025)	(0.687)	(1.466)
Owner attributes					
Homeowner	0.077				
	(0.392)				
Credit grade (AA)					
A	0.750**	0.567*	0.249		
	(2.533)	(1.698)	(0.476)		
B	1.730***	1.211***	1.025*		
	(5.012)	(3.322)	(1.684)		
C	3.131***	2.240***	2.454***		
	(7.861)	(5.473)	(3.983)		
D	3.915***	3.025***	2.893***		
	(8.485)	(6.643)	(4.065)		
E	6.118***	4.998***	5.318***		
	(10.756)	(8.127)	(6.153)		
HR	6.309***	5.070***	5.931***		
	(10.142)	(7.644)	(6.205)		
Delinquencies	0.131				
	(0.604)				
Judgements	0.375			0.707	0.068
	(1.586)			(2.450)	(0.169)
Employment_Status(full time)					
Part-time	-0.611				
	(-0.792)				
Retired	-0.112				
	(-0.185)				
Self-employed	0.364				
	(1.605)				
Income_Range ($ 0 – unable to verify)					
$1 - $24,999	-0.654	-0.691	-0.213	-0.284	-0.455
	(-1.293)	(-1.160)	(-0.243)	(-0.458)	(-0.493)

Table 5-5 continues

	(1) General All	(2) Existing firms Only	(3) New firms only	(4) Existing firms only (no credit grade)	(5) New firms Only (no credit grade)
$25,000 - $49,999	-1.197***	-1.404***	-0.306*	-1.148**	-0.774*
	(-2.748)	(-2.949)	(-0.395)	(-2.328)	(-0.939)
$50,000 - $74,999	-1.184***	-1.282***	-0.176	-0.976*	-0.299
	(-2.673)	(-2.639)	(-0.230)	(-1.939)	(-0.369)
$75,000 - $99,999	-1.378***	-1.215**	-0.161*	-1.026*	-0.009*
	(-2.905)	(-2.340)	(-0.195)	(-1.895)	(-0.010)
$100 000 plus	-1.483***	-1.515***	-0.776*	-1.599***	-1.068*
	(-3.129)	(-2.970)	(-0.909)	(-2.990)	(-1.154)
Firm Attributes					
Existing firms	-0.064				
	(-0.326)				
Information Attributes					
Repeat loans	-0.404	-0.430	-0.748	-0.233	-0.475
	(-2.117)	(-1.892)	(-2.304)	(-0.986)	(-1.386)
Include_picture	-0.456**	-0.255	-0.928**	-0.395	-0.964**
	(-2.105)	(-0.975)	(-2.432)	(-1.452)	(-2.374)
Elaboration	0.194				
	(0.316)				
Loan Attributes					
Requested amount ($1000)	0.021	0.018	0.022	0.011	0.019
	(0.404)	(0.213)	(0.641)	(0.022)	(0.032)
SQRequested amount	0.003				
	(1.267)				
Offer interest rate (%)	0.539***	0.619***	0.641***	0.714***	0.721***
	(9.524)	(10.343)	(7.096)	(11.802)	(8.582)
Time fixed effects	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes
Regional effects	Yes	Yes	Yes	Yes	Yes
Number of Observations	1,390	1,001	390	1,001	390
Pseudo R^2	0.291	0.287	0.337	0.275	0.320
Log Likelihood	-3546	-2564	-934.9	-2605	-958.6

Table 5-6: Factors driving cost of credit on Prosper – with selection effects

This table reports the 2 stage estimates for factors driving cost of credit. In all regressions, the dependent variable is the interest rate (less prime rate). The explanatory variables include owner, firm, and information attributes: credit grade, home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan. Model 1 includes all explanatory variables, whilst Model 2 excludes the credit grade variable.

The controls are requested loan size, offer interest rate, employment status, and income. Time, Region and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

Outcome: *Interest rate* $=\alpha+ \beta_1 \, Owner_i + \beta_2 \, Firm_i + \beta_3 \, Information + \beta_4 \, LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i + u_i$

Selection: *Pr (Approval)* $=\alpha+\beta_1 Owneri+\beta_2 Firmi+\beta_3 InformationAttributesi+\beta_4 LoanAttributes_i +\beta_5 Industryi + \beta_6 Macroi + \beta_7 \, Regioni + \mathcal{E}$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively. t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	(1) 2 stage estimation model **Include** credit grade		(2) 2 stage estimation **excluding** credit grade	
	Outcome Interest rate	**Selection** Pr (Approval)	Outcome borrower_rate	**Selection** Granted
Constant	1.687	-0.711	2.533	-0.510
	(0.687)	(-0.725)	(1.000)	(-0.514)
Owner Attributes				
Home_owner	0.224	0.131***	-0.068*	0.121**
	(0.995)	(2.701)	(-0.286)	(2.487)
Credit_grade (AA)				
A	0.252	-0.173*		-0.176**
	(0.755)	(-1.950)		(-2.175)
B	0.492	-0.306***		-0.290***
	(1.359)	(-3.314)		(-3.312)
C	0.971**	-0.484***		-0.524***
	(2.315)	(-4.859)		(-5.426)
D	2.072***	-0.650***		-0.760***
	(4.555)	(-5.833)		(-6.987)
E	4.688***	-1.082***		-1.389***
	(8.071)	(-8.312)		(-10.868)
HR	3.371***	-1.527***		-1.748***
	(5.083)	(-11.339)		(-13.255)
Delinquencies	0.321	-0.147***	0.262*	-0.168***
	(1.324)	(-3.082)	(1.030)	(-3.270)
Judgements	0.325	-0.203		-0.219
	(0.430)	(-1.543)		(-1.510)
Employment_Status(full time)				
Part-time	-1.978	-0.399**	-2.349	-0.419**
	(-2.122)	(-2.223)	(-2.385)	(-2.310)
Retired	0.573	-0.020	0.771	-0.050
	(0.837)	(-0.146)	(1.066)	(-0.360)
Self-employed	-0.060	-0.088*	-0.235	-0.104*
	(-0.237)	(-1.649)	(-0.877)	(-1.931)

Table 5-6 continues

	(1) 2 stage estimation model **Include** credit grade		(2) 2 stage estimation **excluding** credit grade	
Income_Range ($ 0 – unable to verify)				
$1 - $24,999	0.346	0.207*	0.604	0.189*
	(0.628)	(1.921)	(1.036)	(1.736)
$25,000 - $49,999	0.158	0.270***	0.494	0.273***
	(0.335)	(2.868)	(0.990)	(2.871)
$50,000 - $74,999	0.508	0.321***	0.963	0.331***
	(1.063)	(3.348)	(1.913)	(3.420)
$75,000 - $99,999	0.184	0.251**	0.679	0.270**
	(0.355)	(2.316)	(1.240)	(2.474)
$100 000 plus	-0.208	0.151	0.122	0.158
	(-0.406)	(1.409)	(0.226)	(1.467)
Firm Attributes				
Existing firms	-0.112	0.049	-0.112	0.050
	(-0.461)	(1.040)	(-0.461)	(0.972)
Information Attributes				
Repeat loans	0.121		0.295	0.121
	(0.599)		(1.413)	(0.599)
Include_picture	-1.276***	0.100*	-1.744***	0.052*
	(-5.967)	(1.893)	(-7.831)	(0.991)
Elaboration	-0.455	0.072	-0.551	0.019
	(-0.422)	(0.297)	(-0.481)	(0.078)
Bid_count		0.008***		0.008***
		(35.311)		(34.509)
Loan Attributes				
Requested amount ($1000)	0.005	-0.183***	0.059	-0.237***
	(0.261)	(-28.378)	(0.990)	(-16.039)
SQRequested amount			-0.000	0.002***
			(-0.139)	(3.971)
Offer interest rate (%)	0.719***	0.009***	0.796***	0.021*
	(45.451)	(2.796)	(62.506)	(1.346)
SQ Offer interest rate			-0.000	-0.000
			(-0.464)	(-0.464)
Time fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Regional effects	Yes	-	Yes	-
Number of Observations	10,303	10,303	10,303	10,303
Log Likelihood	-4976	-4976	-5006	-5006
LR test of independence	chi2(1) = 3.13; p<0.001		chi2(1) = 3.06 ; p < 0.001	

Table 5-7: Factors driving cost of credit on Prosper – with selection effects and accounting for ex post default

This table reports the 2 stage estimates for factors driving cost of credit. In all regressions, the dependent variable is the interest rate (less prime rate). The explanatory variables include owner, firm, and information attributes: credit grade, home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan. Model 1 includes all explanatory variables, whilst Model 2 excludes the credit grade variable.

The controls are requested loan size, offer interest rate, employment status, and income. Time, Region and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

Outcome: *Interest rate* $=\alpha+\ \beta_1\ Owner_i\ +\ \beta_2\ Firm_i\ +\ \beta_3\ Information\ +\ \beta_4\ LoanAttributest_i\ +\ \beta_5 Industry_i\ +\ \beta_6 Macro_i\ +\ \beta_7 ex\_default_{i\ +}\ u_i$
Selection: *Pr (Approval)* $=\alpha+\beta_1 Owneri+\beta_2 Firmi+\beta_3 InformationAttributesi+\beta_4 LoanAttributes_i\ +\beta_5 Industryi\ +\ \beta_6 Macroi\ +\ \beta_7\ Regioni\ +\ \mathcal{E}$

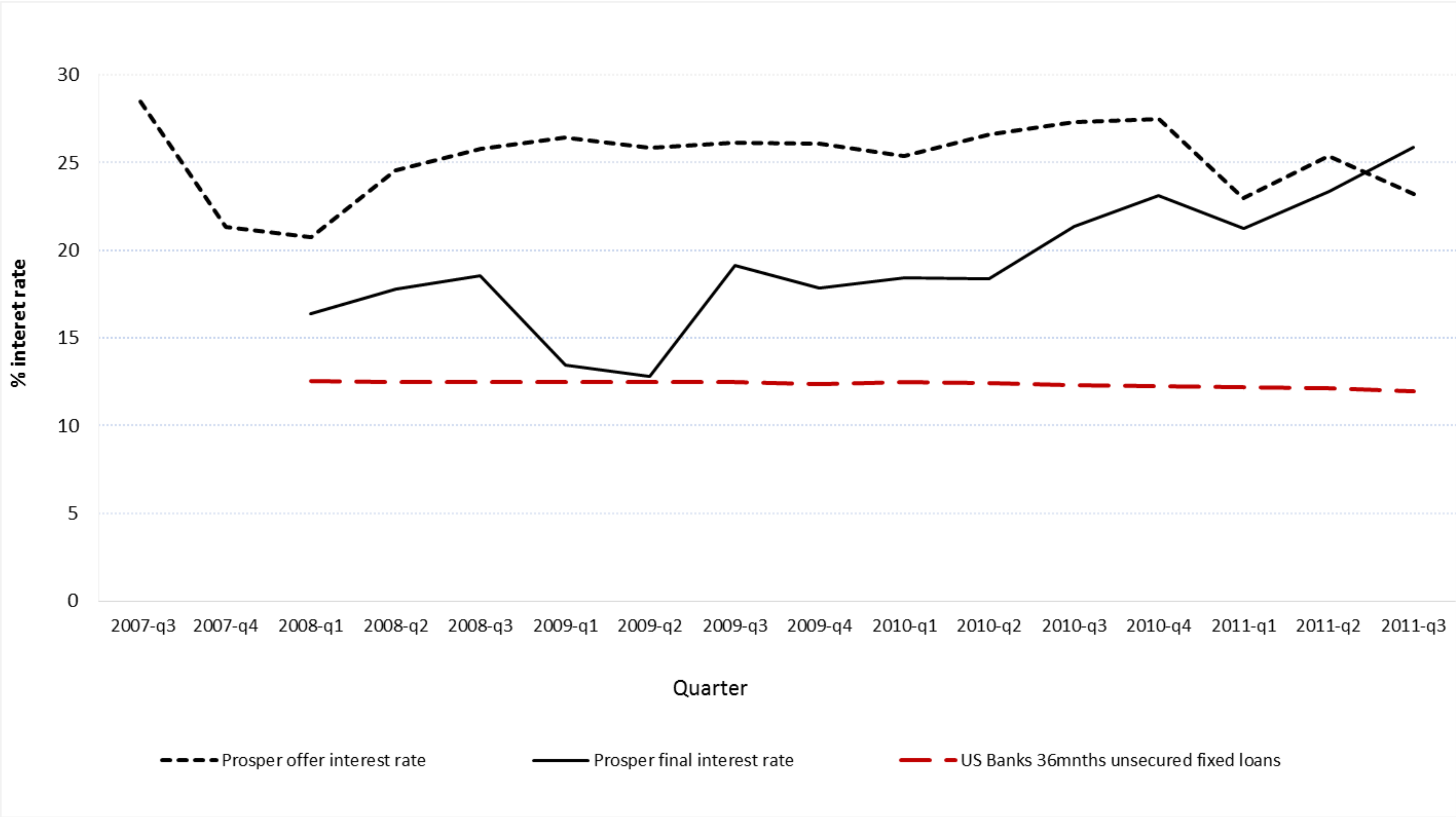
Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively. t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	(1) 2 stage estimation model **Include** credit grade		(2) 2 stage estimation **excluding** credit grade	
	Outcome Interest rate	**Selection** Pr (Approval)	**Outcome** borrower_rate	**Selection** Granted
Constant	1.687	-0.711	2.533	-0.510
	(0.687)	(-0.725)	(1.000)	(-0.514)
ex_default	1.171***		1.687**	
	(4.219)		(0.687)	
Owner Attributes				
Home_owner	0.224	0.131***	-0.068*	0.121**
	(0.995)	(2.701)	(-0.286)	(2.487)
Credit_grade (AA)				
A	0.252	-0.173*		-0.176**
	(0.755)	(-1.950)		(-2.175)
B	0.492	-0.306***		-0.290***
	(1.359)	(-3.314)		(-3.312)
C	0.971**	-0.484***		-0.524***
	(2.315)	(-4.859)		(-5.426)
D	2.072***	-0.650***		-0.760***
	(4.555)	(-5.833)		(-6.987)
E	4.688***	-1.082***		-1.389***
	(8.071)	(-8.312)		(-10.868)
HR	3.371***	-1.527***		-1.748***
	(5.083)	(-11.339)		(-13.255)
Delinquencies	0.321	-0.147***	0.262*	-0.168***
	(1.324)	(-3.082)	(1.030)	(-3.270)
Judgements	0.325	-0.203		-0.219
	(0.430)	(-1.543)		(-1.510)
Employment_Status(full time)				
Part-time	-1.978	-0.399**	-2.349	-0.419**
	(-2.122)	(-2.223)	(-2.385)	(-2.310)
Retired	0.573	-0.020	0.771	-0.050
	(0.837)	(-0.146)	(1.066)	(-0.360)

Table 5-7 continues

	(1) 2 stage estimation model **Include** credit grade		(2) 2 stage estimation **excluding** credit grade	
Self-employed	-0.060	-0.088*	-0.235	-0.104*
	(-0.237)	(-1.649)	(-0.877)	(-1.931)
Income_Range ($ 0 – unable to verify)				
$1 - $24,999	0.346	0.207*	0.604	0.189*
	(0.628)	(1.921)	(1.036)	(1.736)
$25,000 - $49,999	0.158	0.270***	0.494	0.273***
	(0.335)	(2.868)	(0.990)	(2.871)
$50,000 - $74,999	0.508	0.321***	0.963	0.331***
	(1.063)	(3.348)	(1.913)	(3.420)
$75,000 - $99,999	0.184	0.251**	0.679	0.270**
	(0.355)	(2.316)	(1.240)	(2.474)
$100 000 plus	-0.208	0.151	0.122	0.158
	(-0.406)	(1.409)	(0.226)	(1.467)
Firm Attributes				
Existing firms	-0.112	0.049	-0.112	0.050
	(-0.461)	(1.040)	(-0.461)	(0.972)
Information Attributes				
Repeat loans	0.121		0.295	0.121
	(0.599)		(1.413)	(0.599)
Include_picture	-1.276***	0.100*	-1.744***	0.052*
	(-5.967)	(1.893)	(-7.831)	(0.991)
Elaboration	-0.455	0.072	-0.551	0.019
	(-0.422)	(0.297)	(-0.481)	(0.078)
Bid_count		0.008***		0.008***
		(35.311)		(34.509)
Loan Attributes				
Requested amount ($1000)	0.005	-0.183***	0.059	-0.237***
	(0.261)	(-28.378)	(0.990)	(-16.039)
SQRequested amount			-0.000	0.002***
			(-0.139)	(3.971)
Offer interest rate (%)	0.719***	0.009***	0.796***	0.021*
	(45.451)	(2.796)	(62.506)	(1.346)
SQ Offer interest rate			-0.000	-0.000
			(-0.464)	(-0.464)
Time fixed effects	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes
Regional effects	Yes	-	Yes	-
Number of Observations	10,303	10,303	10,303	10,303
Log Likelihood	-4976	-4976	-5006	-5006
LR test of independence	chi2(1) = 2.17 ; p< 0.001		chi2(1) = 2.22 ; p<0.001	

Figure 5-1: Final interest rates paid over time



Source: US Banks 36 months unsecured fixed loan data

Chapter 6 Factors driving small business loan default in P2P lending

Is it worth it for lenders to invest?

6.0 Introduction

Finally, in this study we are interested in understanding the general performance of the funded loans. Looking at the default activity in comparison to lender returns, will help us ascertain whether it is worth it for strangers to invest in small business ventures in this market. The two previous empirical studies put forward some of the mechanisms which lenders seem to adopt in when making credit extension decisions. In this empirical study we investigate who is likely to default. We are especially interested in testing whether by not incorporating firm level characteristics in the credit extension and in the pricing of credit P2P lenders are making better or worse off retunes. Hence we test the following hypotheses:

H1c small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are less likely to default

H2c small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are less likely to default

H3c small business borrowers who have successfully paid back a previous loan are less likely to default on future loans

H4c existing firms are less likely to default (relative to new business start-ups) in the P2P the lending context

H6c small business borrowers who use text elaborations are more (less) likely to be default

H7c small business borrowers who have previous failures are more likely to default

H8c small business owners who post pictures are less (more) likely to default

6.1. Data

The analysis is based on loans funded between 2007 and 2011. Prosper provides information on payment status for each loan. The loans are annuities, with repayments in equal monthly instalments. Repayments are automatically deduct from the borrower's account and distributed to lenders' Prosper accounts. If the monthly payment is made on time, the loan status is considered current. If a monthly bill is not paid, the loan status will be changed to 1 month late, 2 months late and so on. If a loan is late for more than 90 days, it is considered to be in default. All defaulted loans are reported to credit reporting agencies and can affect borrowers' credit scores. Borrowers who default on their loans are not permitted to borrow using Prosper again.

For the purpose of this analysis, we consider a default to have occurred when the principal loan amount (plus the calculated interest) remains unpaid in full when the loan reaches the maturity date. Our initial sample consisted of 1417 funded loans of which 1208 (87 percent) have reached full maturity - 885 loans (64 percent) were paid off; 323 loans (23 percent) resulted in default; while 209 loans (15 percent) are current. We then apply the following filtering criteria: Loans initiated in 2011 require complete payment information in 2014 for us to judge whether or not a default has occurred. In total 209 loans found in our sample were granted in 2011; 177 are registered as current, 59 loans have reached full maturity and were paid off and 39 loans were in default. Including only the sub sample of defaulted loans from the 2011 cohort may bias our estimates, i.e., we may either under or overestimate default. Consequently, for consistency with our default definition, we remove 209 such loans. The final sample therefore consists of 1208 funded loans, 302 (25 percent) of which resulted in a default.

6.2 Univariate analysis

Table 6-1 present descriptive statistics of the default study sample. In Table 6-1 there are 4 columns: column 1 shows statistics for all funded loans while columns 2 and 3 show statistics for paid loans and those which resulted in default respectively. Of the 1208 funded loans, 25 percent have resulted in default. A simple comparison of means between paid loans and defaulted loans, found in column 4, give a basic intuition on factors that might be associated with default. Variables showing statistically significant mean differences between paid and defaulted loans include *Credit_grade*, *Home_ownership*, *Repeat_loans* and *Include_picture*.

...........Table 6-1 goes around here...........

Of the significant variables, if a small business borrower is a home owner, this reduces the proportionality that they will default on the loan. The proportion of homeowners for borrowers that repaid their loans is 57 percent compared with a 59 percent of home owners who failed to repay their loans ($p<0.01$). Similarly, if a small business borrower is in a premium credit grade (eg AA), this reduces the proportionality that they will result in default (29 percent repay loans vs. 14 percent that default); whilst if the borrower is in a high risk category this heightens the proportionality that they will result in default (6 percent vs. 3 percent).

Turning to information variables, being a repeat borrower reduces the proportion that the loan will result in default. The proportion of repeat loans for small business borrowers that repaid their loans is 43 percent (compared to 22 percent that defaulted). Similarly, including a picture seems to reduce the proportion that the loan will result in default (61 percent vs. 69 percent).

The remaining explanatory variables namely: *Delinquencies*, *Judgements*, *Existing_firms*, and *Elaboration* did not show statistically significant mean differences between the two cohorts. Moreover, we see from Table 6-1 that small business borrowers that defaulted on their loans on average asked for larger loan amounts ($9,681 vs. $7507); and on average they typically paid higher interest rates than those who repaid the loans (23 percent vs. 17 percent).

We observe from Table 6-2 that 27 percent of all defaulted loans resulted within the first year of origination. This latter measure, to some extent, may indicate the soundness of the lenders' decision making when appraising credit risk on Prosper. By the end of the second year, well over 70 percent of all defaults have occurred. This may be an early indication that high numbers of loan default are the rule rather than the exception in this market, perhaps indicating lender's inexperience.

...........Table 6-2 goes around here...........

Figure 6-1 shows how the observed default rates have changed over time. There was a dip in default rates between 2008 and 2010; but after the first quarter of 2010, default rates showed a linear increase for the remaining duration of our data sample. At least three possible explanations may help us account for this observation: macroeconomic factors, evolution of Prosper's credit risk over time and an increase in interest rate caps enabled by Prosper across 56 US states in April 2008. As shown in Figure 6-1, the unemployment rate increased in 2009 (almost doubled when compared to previous year); subsequently we observe an increase in default towards the end of 2009, a trend that continued throughout peaking in the second quarter of 2011. Furthermore, Figure 6-1 suggests that overall the pool of loans has worsened over time on Prosper. Although loans initially migrated away from the worst credit grades; around the second quarter of 2010 the fractional change has since surpassed early 2008 levels, moving towards worst credit grades - a possible explanation of the observed upward movement of default rates over time.

...........Figure 6-1 goes around here...........

Overall, given the observed difference in mean value for some of the variables of interest between observed for defaulted and non-defaulted borrowers; an examination of the mean values provides motivation for the regression analysis which follows, as the two groups have some observed differences.

6.3 Regressions

We now move on to the first regression that estimates the relationship between the three main categories of explanatory variables and the probability that the loan will result in default, whilst controlling for loan attributes as well as industry and macroeconomic factors. The Probit models are estimated based on equation 6-1 below. In other words, we model the likelihood that '*Default=1'*. All explanatory variables (are defined in Table 3-1 in Chapter 3). Furthermore, as shown in the bivariate correlation matrix (in Table 3-2 in Chapter 3), in terms of explanatory variables, where significant, none of the correlations are above 0.5 demonstrating that in relation to the econometrics models, multicollinearity is unlikely to be a problem.

$$Pr\ (Defaultl|1) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information\ Attributes + \beta_4\ Loan\ Attributes + \beta_5 Industry_i + \beta_6 Macro_i + \mu \qquad (6\text{-}1)$$

The results are shown in Table 6-3. As stated previously, in all the estimations, a general- to - specific procedure to establish the model of interest was adopted (Darlington, 1990). Column 1 of Table 6-3 shows the probit estimates for the general specification; which includes all the regressors, identified from literature, which could possibly be important determinants of default in the P2P lending context. The procedure begins with a full model and removes variables with a p-value above 0.10 resulting in a parsimonious model as shown in column 2. The results of the Likelihood Ration (LR) test confirm that the parsimonious model is a better estimation; the null hypothesis that the excluded regressors collectively have no role in predicting our dependent variable is decisively accepted (LRX^2= - 511, *df* = 102, p <0.10). The third and fourth columns subsequently report the determinants of credit approval separate for existing firms and new business start-ups - with the main aim of testing the proposition firmly established in literature that new business start-up are more likely to default relative to already established firms.

...........Table 6-3 goes around here...........

In general, our results show that default is found to be related to risk as predicted by conventional theory in both formal credit rating and also information and track record (see for example Berger *et al*, 2005; Agarwal *et al*, 2007; DeYoung *et al*, 2007; Berger *et al*, 2009).

More specifically, consistent with what we have already seen in the univariate analysis, the coefficients of the *Credit_grade* variable are positive and statistically significant. This indicates that lenders on Prosper, similar to traditional lenders like banks, do indeed lend with a conservative mind set - small business borrowers in high risk credit grade categories are more likely to default (DeYoung *et al*, 2007; Berger *et al*, 2009). We notice in fact from column (2) that the coefficients of the credit grade variable gets progressively stronger as we move from low risk borrowers (A) to high risk borrowers (HR) in predicting default risk (H_{1c}). This result underpins the importance of credit scoring as a method of predicting default, supporting results put forward by conventional theory (Berger et al, 2009; Berger and Frame, 2007). In terms of the remainder of owner variables - we find that *Home_ownership*, *Delinquencies,* and *Judgements* are not predictors of default in this context. This result is persisting even in the absence of the *Credit_grade* variable (given that we have shown *Credit_grade* can incorporate some of the information already captured by these other variables). Hence H_{1c} and H_{7c} are not supported. One interpretation which can be offered for this observation is that perhaps this finding highlights characteristics of the P2P lending market – when lenders default, the borrower's home cannot be taken to compensate the lenders. Loans are unsecured.

Importantly, in terms of our firm variables, we find that new small business start-ups are no more likely to result in default when compared to established firms our variable of interest, *Existing_Firm* is not statistically significant. This result is counter to what is typically found in banking literature where new firms are perceived to be the riskier class on the bases of severe information asymmetries (Stiglitz and Weiss, 1981, Cassar, 2004). Hence we do not find support for H_{4c}. We also observe that default activity does not favour any industry or sector – all industry dummies were not statistically significant. Therefore, from these results we assert that firm level information is not an important determinant of default in the P2P lending context.

In terms of the information variables - reputation built over time through re-payment of loans attained from prosper decreases the likelihood of default ($p<0.001$). This result seems to offer support to Diamond (1989); affirming that information accumulated specific from Prosper is useful for lenders in determining default risk; especially given the fact that there are no opportunities to collect information through physical interactions (H_{3c}). Hence, our findings implies that a presence of a track record (developed on Prosper) matters and is useful in predicting default in the P2P lending context. Interestingly, we also find that reducing information gaps improves default activity (Stiglitz and Weiss, 1981; Bester, 1985). All else equal, we find that small business borrowers that include a picture in their loan request are less likely to default. The variable *Include_picture* is negative and significant at the 0.01 level; suggesting perhaps that inclusion of a picture somehow either humanizes the process or gives some information about the business or product which lenders seem to use to separate credit risk (H_{8c}). Seemingly, the stories that small borrowers tell may not necessarily influence default; the variable Elaborate is statistically insignificant (H_{6c}).

Looking at the control variables, a key finding of our result is that the cost of credit is associated with default. All else equal, we find that those high interest rates translate to a higher likelihood of default; the variable *Final_interest_rate* is positive and significant at the 0.01 level. This result seems to support the moral hazard argument developed by the model of Stiglitz and Weiss, (1981) where small business borrows (in the absence of collateral) may be encouraged to take additional risk. After all, if their projects succeed, they will keep all the gains - and yet when they fail, it is the P2P lenders that will lose their capital (Stiglitz and Weiss, 1981). This finding presents an interesting dilemma for lenders. For example, if holding all else constant, and lenders opt to lower interest rates that they will accept from prospective borrowers on a loan, and this reduces the default probability, at what point do lenders make a higher expected return even with lower income stream? This argument seems to suggest that there may be a potential or real trade-off between interest rates and default which lenders might have to take into considerations; which presents an interesting dilemma for lenders. However, it is also plausible that some lenders extend credit to firms in this context for fun - as some form of 'gambling'. Hence it would not be necessary to consider a potential trade-off between interest rates and default. In this case, lenders would simply accept loan requests from borrowers offering high interest rates with the view that knowing that if the borrowers pay back the loan, they win big. However, if some of the loans in the portfolio result in default, the loss is not too big. Furthermore, it is also reasonable to consider that some of the lenders choose to extend credit in this context driven by philanthropic reasons. This is supported by evidence we found earlier that lenders in this market extend credit focusing on people – hence their idiosyncrasies may be at play. Therefore, lenders driven by philanthropy might not necessarily have wealth maximisation as his main goal for funding these firms (see for example Argawa et al, 2011).

Finally, our results show that larger loans are more likely to result in default, as evidenced in the positive sign and high significant coefficient of the variable *Requested_amount.* This highlights the acute moral hazard issues entrenched in P2P lending. Typical in traditional lending; large credits usually induce higher borrower motivation because the borrower stands to lose a lot in the event of default given that borrowers offer collateral (Hanley and Crook, 2005). But in this case, since these loans have no security – it seems moral hazard risk is heightened. We can also interpret this result in the context of borrower motivation: smaller credits are taken very seriously in P2P lending. All things equal, the cost of default - which in this case default may result in a judgement and the threat of losing access to an external source of credit - may be damaging to a small business borrower (especially those in the early stages) who might otherwise have limited avenues of external (unsecured) small business finance.

Next, in order to confirm the result as shown by the indicator variable, *Existing_firm*, in column 3 and column 4 of Table 6-3 we extend the analysis to determine factors associated with the likelihood of default separately for existing firms and new business start-ups as defined by equations (6-2) and (6-3).

$$Pr\ (default{:}\ Existing\_firms|1) = \alpha + \beta_1 Owner_i + \beta_2 Information + \beta_3 LoanAttributest_i + \beta_4 Industry_i + \beta_5 Macro_i + \mu \quad (6\text{-}2)$$

$$Pr\ (default{:}\ New\_business\_start\text{-}ups|1) = \alpha + \beta_1 Owner_i + \beta_2 Information + \beta_3 LoanAttributest_i + \beta_4 Industry_i + \beta_5 Macro_i + \mu \quad (6\text{-}3)$$

We see from Table 6-3 that the two key variables that help disentangle information borrower risk namely, the *Credit_grade* variable and the *Include_picture* variables differ between new business start-ups and established firms. All other factors driving the likelihood of default for new firms and established firms are generally similar. More specifically, in both cases the determinants of default include the following variables: *Repeat_loan, Requested_amount, Offer interest_ rate.* Interestingly, we find that although *Credit_grade* remains significant for established firms; the variable is insignificant in the new business start-ups estimation. This result seem to suggests that, in the absence of a track record, lenders may be unable to disentangle credit risk associated with small business borrowers starting new firms. Interestingly, we also see that, inclusion of a picture becomes insignificant for the established firm, and remains negative and significant for new business start-ups. One explanation for the modest role for *Include_picture* in determining default for existing firms could be that perhaps better borrowers have learnt that lenders in the market simply values pictures in general and they tend to include pictures. Finally, looking at loan characteristics, for both new business start-ups and established firms, we find that loans with higher interest rates are indeed more difficult to repay - lending support to the fact that higher interest rates are an indicator of higher risk. Similarly, borrowers with larger loans are more likely to drive default.

6.4 Average marginal effects

To assess the impact of the observed factors, that drive default risk, we compute marginal effects at the mean as shown in Table 6-4. For an otherwise average small business owner with a C credit rating, all else equal, the predicted probability of default is 8 percentage points greater when compared to the prime credit grade AA.

...........Table 6-4 goes around here...........

When this business owner increases the loan amount by $1,000 (from $8,108 to $9,108) we observe an increase in default risk of 1.3 percentage points; meaning a 5 percent increase in default probability. Increasing the requested loan amount by a factor of 10 however, to $18,108, the default probability increases by 13 percentage points; meaning a 52 percent increase in default probability- which may render the loan unplayable.

Likewise, we see from Table 6-4 that increasing the interest rate by 1 percentage point (for example from 18 percent to 19 percent), meaning a 5 percent increase in default probability. Similarly increasing the interest rate by a factor of 10 from 18.1percent to 28.1percent, results in a 52 percent increase in default probability.

We also observe find that all else equal, the predicted probability of default is decreased by 7.7 percentage points if the small business owner includes a picture. In relation to the average probability of default, 25 percent, a difference of 7.7 percentage points means a 31 percent decrease in the likelihood that the loan will result in non-payment. Interestingly, the impact of pictures is especially substantial for new firms, such that including a picture results in a decrease in default probability of 15.7 percentage points. Compared to an average probability of default, 25 percent, this represents an approximately 62 percent in reduction in the likelihood of default.

6.5 Margins at representative values

Table 6-5 puts into perspective our argument above – showing how the default rate varies at representative values. We observe from the table that simply improving on the credit grade, reduces the default probability substantially – all else equal a borrower with an AA credit grade has a 10 percent chance of defaulting on the loan; whilst a high risk borrower with similar characteristics will almost have a 60 percent chance of defaulting when compared to a borrower with an average credit grade.

...........Table 6-5 goes around here...........

6.6 Robustness Check

Our probit model estimate the default equation based on a sample of funded loans, selected non-randomly from the sample of all loans (funded and not funded). If selectivity exists; the coefficients we observe from the probit estimate may not be applicable to all borrowers requesting loans. To determine whether selection is a problem, we run a robustness check - by estimating a 2 stage Heckman selection model. The first stage of the Heckman's model estimates a probit in order to determines whether the loan was funded or not; and the second stage estimates a probit model to determines factors which drives default (given that the loan was funded).

We estimate the first stage as a function of the original control variables credit used in the default model above (i.e. credit score, collateral, firm age, loan size, interest rate, information, and controlling for industry, region and time effects) and an additional identifying variable - in this case the borrower income range. Bank literature suggests that borrower income is expected to affect the probability of a borrower being funded positively; it determines the affordability of the borrower to be able to pay back the loan. While being relevant for funding decision, borrower income range should not influence the default equation. Once lenders have appraised the borrower's affordability of loan repayment and compensated with an appropriate risk premium, lenders should be indifferent with respect to borrower income range. The estimation results of the Heckman equation are reported in Table 6-6. Indeed when we look at estimation result in column 2, income rage has a positive and significant effect on the probability of funding; we observe that small business owners that have an income are more likely to be granted a loan than those in the zero or undefined income range category. Income range is insignificant in the default equation (column 1 of Table 6.6).

Stage 2 Outcome regression: observed only if loan funded (i.e. Pr *(Approval)* = 1)

$$Pr\ (Default) = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ InformationAttributes_i + \beta_4\ LoanAttributes_i$$

$$+\beta_5 Industry_i + \beta_6 Macro_i + \beta_7\, Region_i \;\; + \mu \qquad \text{(6-4a)}$$

Stage 1 selection regression: probability of credit approval is 1 if loan funded, and 0 otherwise

$$Pr\ (Approval) = \alpha + \beta_1 Owner_i + \beta_2 Firm_i + \beta_3 InformationAttributes_i + \beta_4 LoanAttributes_i$$
$$+\beta_5 Industry_i + \beta_6 Macro_i + \beta_7\, Region_i + \mathcal{E} \qquad \text{(6.4b)}$$

The LR test of independence (X^2 = 1.16; p=0.001); which confirms the adoption of the Heckman estimation to control for selection bias. In general, the coefficient estimates, signs and statistical significance for our entire variable set are maintained in the two-stage regression. Hence, our robustness checks confirm the earlier obtained result.

...........Table 6-6 goes around here...........

Our results for new and established businesses in Probit regressions seem to suggest that Firm status does not necessarily drive default. The Heckman result, however, seem to suggest we exercise a little caution in this interpretation – especially given that after accounting for selection, credit grade variable is insignificant for new business borrowers. This could suggest that it may be difficult to appraise risk for this cohort.

6.7. Estimation of Net Returns

In this section we examine the net returns that lenders get from investing in P2P small business loans, with the view of gaining insight as to whether it is worth it for lenders to invest through this medium.

In Figure 6-3 we show average lender returns on Prosper, over time. We find that mean return for P2P small business loans is 3.26 percent per year (having charged on average 18.5 percent interest for the loan that has a 25 percent probability of default). Returns in this market are somewhat volatile, with an annual standard deviation of 3.67 (i.e. over 100 percent). In Table 6-7 we illustrate the proportion of total lending attributed to the top 5 percent of lenders; based on the total dollar amount lent. The single biggest lender has disseminated $125, 000 worth of small business loans across 23 loans; which amounts to 0.8% of the total lending for the whole sample (i.e. derived as follows $125thousand/ $15.4mil). We observe from Table 6-7 a risk adjusted average return of 6.1 percent per annum for the top 5 percent of lenders (having charged an average interest rate of 18.5 percent) with a less erratic annual standard deviation of 1.24 (i.e. 20 percent).

...........Table 6-7 goes around here...........

From these results we can make three key assertions: First, we can make an assertion regarding the risk tolerance of the individual lenders in this market. On average, the standard deviation associated with the returns of loans from the top 5 percent lenders was relatively low. By contrast, for the average lender pool, the standard deviation of returns associated with those loans was high. As a result, based on standard deviation alone, we might come to the conclusion that the average lender on P2P platforms actually has a higher risk tolerance (a different class of investors to those inferred in the model by Stiglitz and Weiss, 1981) . Hence, it can be inferred then that the utility function of the average lender on Prosper is that of an upward slope. For these lenders, the risk of losing a small proportion (as little as $25) per investment in the overall portfolio of loans is offset by the potential gain from high interest rates charged for loans. Consequently, given the fact that lenders in this market are risk lovers, perhaps the 1.49 percent return from a less risky or more conservative bank investment and a 1.61 percent return from a less risky Treasury bond may seem less appealing to this cohort. These interest rates are low; investors are possibly looking for better returns.

Second, we can further make an assertion on how informed (or uninformed) the lenders are. By comparison, the return achieved by an average lender on Prosper is 3 percentage points lower when compared to the returns achieved by the top 5 percent lenders (i.e. 3.26 percent returns vs. 6.1 percent returns). Our result seems to confirm our initial observation of the fact that the average lenders on Prosper are indeed amateurs in making investment decisions – different class of investors to that inferred in the model of Stiglitz and Weiss (1981). Only a handful of lenders are able to attract returns almost comparable to other classes of risky assets such as returns from the stock market or returns from 3 year venture capital investments (6.1 percent vs. 7 percent stock market vs. 8.5 percent[5] from VC). Put differently, our results suggest that P2P lending introduces a new breed of informal investors into the small business finance landscape, who are risk loving enough to look for high yield investment opportunities but who are not conversant enough to be able to participate in more technical asset classes like the stock market.

Third, we can make some assertions concerning the viability of P2P lending in the long run. Given the fact that P2P lending is generally a young market, and the fact that majority of lenders attracted to P2P lending are relatively uninformed amateurs in making investment decisions, we observe from our results that if the amateur lenders do indeed learn, it then becomes plausible that in time the returns in this market may generally converge to be better (and gravitate towards the 6.1 percent achieved by top 5 percent). However, if the P2P lending platforms continue to attract a pool of amateur lenders, the average returns of 3.25 percent may render the market somewhat unsustainable in the long run. Given the current data limitation however, viability of P2P lending remains an avenue that needs to be further explored by future research as the market matures and more performance data become available.

...........Figure 6-3 goes around here...........

So who makes money in this market? We observe from Figure 6-4 that 74 percent of funded loans results in a positive return - on average lenders that make money, make around 9 percent per annum. Just over 20 percent of the loans however result in a complete loss while the remaining 5 percent of the loans either breaks even, or incur some loss.

...........Figure 6-4 goes around here...........

[5] **See Table 5-2;** Data Source: Damodaran (2010): The Dark Side of Valuation: Valuing Young, Distressed, and Complex Business

When we breakdown the characteristics of best performing returns by loan size, Figure 6-5 clearly shows a strong negative skew – with almost 40 percent of the loans less than $5,000; perhaps representing the (lower) risk involved. All else equal, it seems smaller loans are more profitable.

...........*Figure 6-5 goes around here*...........

When we breakdown the characteristics of best performing returns by credit rating as shown in Figure 6-6, as to be expected, the higher risk categories give higher returns to compensate for the relative risk.

...........*Figure 6-6 goes around here*...........

We observe from Figure 6-7 that as the market ages, the distribution of loans changes such that default rates decrease - resulting with better returns. So, we see that lenders that extend credit to borrowers that have no history in the platform, on average they can expect to get 1 percent return. But as firms continue to build history, returns increase.

...........*Figure 6-7 goes around here*...........

Theory stipulates that new firms are riskier (Cassar, 2004). When a lender invests in a new firm, they will receive an average return of 2.5 percent (given a default risk of 26.6 percent with borrowers paying 18.2 percent interest rate). If a lender chooses to invest in an existing business instead, they stand to make on average 3.5 percent (given a 23.7 percent default rate and borrower paying on average 17.8 percent). We see from Figure 6-8 that lenders investing in new businesses that have no pre-established payment history in this market will result with complete losses - start-ups result in losses when they first enter the market; those that choose to invest in existing firms that come to P2P for the first time, will make on average a return of 1.6 percent. We observe however from Figure 6-8 that over time the returns from both new and existing firms converge; and by the 5th loan, lenders' returns average at 10 percent from both new firms and existing (which may be a knock on confirmation that over time start-ups were no more likely to default than existing firms).

...........Figure 6-8 goes around here...........

Finally, perhaps our most interesting result is that of including an image. Lenders who invest in new firms that include a picture stand to make on average 3.7 percent returns; whilst those that invest in new firms that do not include a picture receive on average 0.2 percent return.

6.8 Chapter summary

The overall research objective of this study was to determine whether P2P lending is sustainable in the long run. We intended to gain this knowledge by first understanding key determinants driving default in the P2P lending context, then second, by estimating lender returns. This is one of a smaller number of papers that exploits micro level data from P2P lending websites on defaults among small business borrowers.

We find that small business owner's credit grade, the requested loan size, and interest rate paid, are significant predictors of default in this market. Our results also lend support to the prediction that new firms are no more likely to default when compared to existing firms. This is a particularly important finding as start-ups are widely perceived to be the riskiest class of small firms. We observe that small business borrowers anticipating to start new businesses that include pictures tend to be more successful in reducing information asymmetries, allowing lenders to conduct better due diligence – loans that include picture are less likely to default. We see from the analysis that smaller loans offer better returns. Moreover, loans from riskier borrowers in terms of credit grades offer higher returns to compensate for the relative high risk.

Furthermore, we also report in the study that default does not vary across different types of industries such that P2P lending becomes sustainable only in some parts of the small business sector and not in others – all else equal; industry variables are insignificant determinants of default.

Is it worth it for lenders to extend credit to these firms? The distribution of returns for this type of investment is quite varied. On average lenders will receive 3.26 percent per annum for their investment in small business loans (having charged on average 18.5 percent interest for the loan that has a 25 percent probability of default). Returns in this market are somewhat volatile, with an annual standard deviation of over 100 percent. Our results suggest that average lender on P2P platforms actually has a higher risk tolerance relative to those lenders described by the model of Stiglitz and Weiss (1981), hence, it can be inferred then that the utility function of the average lender on Prosper is that of an upward slope. For these lenders, the risk of losing a small proportion (as little as $25) per investment in the overall portfolio of loans is offset by the potential gain from high interest rates charged for loans.

Overall our results suggest that P2P lending is some kind of 'gambling' for lenders - if a portfolio with relatively riskier business ventures are chosen, the result will exceed the market average (of 3.26 percent) but if not - the loss is not too big. In fact, our results introduce a novel finding, that is P2P lending introduces a new breed of informal investors into the small business finance landscape, who are risk loving enough to look for high yield investment opportunities but who are not conversant enough to be able to participate in more technical asset classes like the stock market.

We make key contributions. First, we shed light on the determinants of default in the P2P lending context; highlighting that borrower reputation continues to be the single most dominant determinant of default. Second, we find that the breed of investors extending credit to entrepreneurs in this context are amateurs, who have high risk tolerance, with a completely different utility function to that of lenders inferred in the model by Stiglitz and Weiss (1981). This study sheds some light that P2P lending may be availing a different type of investor to that previously seen in the small business lending literature. Finally, Caution will to be exercised when ascertaining whether our findings can be generalised to developing countries for example. As it stands, P2P lending does not have other mechanisms at its disposal such as peer pressure and a sense of community, which help keep default rates down when extending microfinance. Given that relationships in this context are very weak at best, since lending takes place online, default rates may therefore sky rocket in the context of Microfinance institutions, thus rendering P2P lending an unsustainable form of extending credit to small businesses.

Table 6-1: Univariate table: Descriptive statistics for loan default

In this table column (1) shows descriptive statistics all funded loans. In column (2) we show statistics for all loans which were paid back whilst column (3) shows all the loans that resulted in default. Finally, column (4) presents *t-test/x^2 statistics* for differences in the means of the repaid and defaulted loans. *, **, *** stand for 0.1, 0.05, and 0.01 significance levels respectively

Variable	(1) Funded loans	(2) Default=0	(3) Default=1	(4) *t- test / x^2 statistics*
Number of observations	1417	0.75	0.25	-
Owner attributes				
Home_ owner	0.58	0.59	0.57	0.68
Credit_ grade				
AA	0.25	0.29	0.14	4.71***
A	0.19	0.19	0.19	0.11*
B	0.19	0.19	0.19	0.06*
C	0.16	0.16	0.15	0.24
D	0.13	0.11	0.17	-2.8
E	0.05	0.04	0.09	-3.46
HR	0.04	0.03	0.06	-2.65
Delinquencies	2.3	2.3	3.9	-1.4
Judgements	0.27	0.24	0.37	-2.3
Firm attributes				
Existing _firms	0.73	0.74	0.70	1.06
Industry				
construction	0.01	0.01	0.01	0.61
transport and utilities	0.02	0.02	0.02	-0.01
services	0.40	0.40	0.39	0.25
retail trade	0.29	0.27	0.33	-1.87
finance &real estate	0.21	0.23	0.16	2.16**
agriculture	0.02	0.02	0.02	1.74**
wholesale trade	0.01	0.01	0.01	0.61
manufacturing	0.05	0.04	0.09	-2.17
Information Attributes				
Repeat_loan	0.39	0.43	0.22	6.1***
Include image	0.67	0.69	0.61	2.31***
Elaboration	0.99	0.99	0.98	1.24
Loan Attributes				
Requested amount ($ 1,000)	$8,108	$7,507	$9,681	-4.5
Interest rate (%)	18.51	16.97	22.8	-10.3
Macro and Industry attributes				
Unemployment rate (%)	6.7	6.8	6.6	1.58

Table 6-2: Loan default across time

In this table column (1) shows all funded loans while column (2) shows the percentage of loans that resulted in default. In column (3) we present the percentage of loans which resulted in default within the first year of issue while in column (4) we present the percentage of loans which resulted in default within the second year of issue.

Year	(1) All	(2) D=1(%)	(3) D=1 1st yr. (%)	(4) D =1 2nd yr. (%)
2007	8	25%	0%	0%
2008	692	27%	28%	75%
2009	83	12%	20%	50%
2010	146	21.%	29%	48%
Total	929	25%	27%	75%

D1 – Denotes loans which resulted in default, D0 – Denotes loans which were repaid

Table 6-3: Determinants of loan default for P2P small business loans

This table reports the Probit regression results for factors driving default on Prosper.com. The first two regressions present estimates for the general and parsimonious specifications for all loan requests. The last two regressions present estimates for parsimonious specifications for loans from existing firms only and for loans from new firms only. In all regressions, the dependent variable is binary taking the form 1 if credit the loan request was funded and 0 otherwise. The explanatory variables include owner, firm, and information attributes: credit grade, home ownership, repeat loans, delinquencies in the past 10 years, judgements in the past 10 years, firm age, inclusion of pictures, indication of text elaboration, and number of lenders extending credit per loan.

The controls are requested loan size, offer interest rate, employment status, and income. Time and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

$$Pr\ (Default|1) = \alpha + \beta_1\, Owneri + \beta_2\, Firmi + \beta_3\, Information + \beta_4\, LoanAttributesi + \beta_5 Industryi + \beta_6 Macroi + \mu$$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively.

Variables	(1) General All firms	(2) Reduced All firms	(3) Existing Firms only	(4) New Firms only
Constant	-1.904***	-2.145***	-2.034***	-1.937**
	(-2.740)	(-11.871)	(-3.655)	(-2.436)
Owner attributes				
Credit grade (ref AA)				
A	0.365**	0.339**	0.394**	0.255
	(2.296)	(2.190)	(2.082)	(0.781)
B	0.391**	0.373**	0.438**	0.307
	(2.322)	(2.296)	(2.185)	(0.897)
C	0.382*	0.330*	0.580**	0.169
	(1.952)	(1.742)	(2.424)	(0.453)
D	0.637***	0.567***	0.723***	0.476
	(3.042)	(2.841)	(2.851)	(1.187)
E	0.556**	0.509*	0.384*	0.901
	(2.024)	(1.911)	(1.130)	(1.720)
HR	0.701**	0.600**	0.715*	0.748
	(2.289)	(2.027)	(1.901)	(1.297)
Home_owner	0.068			
	(0.654)			
Delinquencies	0.004			
	(0.037)			
Judgements	0.107			
	(0.854)			
Firm attributes				
Existing_firms	0.038			
	(0.348)			
Information attributes				
Repeat_loan	-0.627***	-0.662***	-0.627***	-0.720***
	(-5.438)	(-5.346)	(-4.523)	(-2.967)
Include_picture	-0.212**	-0.170**	-0.155	-0.288**
	(-2.105)	(-1.753)	(-1.250)	(-1.502)
Elaboration	-0.277		-0.174	-0.724
	(-0.578)		(-0.022)	(-1.097)

Table 6-3 continues

Table 6-3 continues

Variables	(1) General All firms	(2) Reduced All firms	(3) Existing Firms only	(4) New Firms only
Loan attributes				
Requested amount ($1,000)	0.052***	0.052***	0.046***	0.071***
	(6.048)	(6.385)	(4.392)	(4.241)
SQRequested amount	-0.002*			
	(-1.073)			
Final_interest_rate (%)	0.040***	0.043***	0.046***	0.029**
	(5.368)	(5.855)	(5.156)	(1.811)
SQFinal_interest_rate	-0.001			
	(-1.389)			
Unemployment rate	YES	YES	YES	YES
Industry dummies	YES	YES	YES	YES
Regional dummies	YES	YES	YES	YES
Observations	1208	1208	845	363
Pseudo R^2	0.144	0.131	0.156	0.165
X^2	-503.7	-511.5	-355.3	-136.0

Table 6-4: Marginal effects

This table presents the marginal effects after Probit models based on all variables set at their means. The marginal effects for categorical variables show how Pr (default = 1) is predicted to change as a particular factor variable changes from 0 to 1, holding all other independent variables at zero. The marginal effect of a continuous variable measures the instantaneous rate of change, which may or may not be close to the effect on Pr (Default=1) of a one unit increase in the independent variable. *, **, *** stand for 0.1, 0.05, and 0.01 significance levels respectively.

	All firms	Existing firms	New firms
Owner attributes			
Credit grade (ref AA)			
A	0.096	0.118	0.052
	(0.041)**	(0.053)**	(0.097)
B	0.082	0.120	0.104
	(0.043)*	(0.055)**	(0.100)
C	0.088	0.168	0.025
	(0.051)*	(0.066)**	(0.110)
D	0.162	0.206	0.143
	(0.058)***	(0.069)***	(0.119)
E	0.171	0.138	0.280
	(0.080)**	(0.092)*	(0.154)
HR	0.151	0.192	0.067
	(0.090)*	(0.102)*	(0.176)
Homeowner	0.023	0.009	0.010
	(0.028)	(0.033)	(0.062)
Firm risk			
Existing_firm	0.010		
	(0.029)		
Information attributes			
Repeat_loan	-0.164	-0.155	-0.294
	(0.028)**	(0.036)***	(0.073)***
Include picture	-0.077	-0.045	-0.157
	(0.030)**	(0.035)	(0.060)***
Elaboration	-0.156	-0.004	-0.189
	(0.148)	(0.203)	(0.171)
Loan characteristics			
Requested amount ($1,000)	0.013	0.013	0.017
	(0.002)***	(0.003)***	(0.005)***
Interest rate (%)	0.013	0.013	0.012
	(0.002)***	(0.002)***	(0.005)**
N	1208	845	363

Table 6-5: Predicted margins at represented values

In this table we present predicted probabilities of loan default at represented values conducted after the Probit regression; with actual cases extracted from the population of loan requests.

Actual Cases	Probability of funding	Interest rate	Default rate
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, homeowner, existing firm, image, elaborate	20%	**12.5%**	10.8%
Full-time employed, Prime credit grade **AA**, income range $50k - $74,999k, no past due loans and no judgments, rent home, existing firm, no image , elaboration	15%	**13.0%**	10.2%
Self-employed, Prime credit grade **AA**, income range $25k - $49,999k, no past due loans and no judgments, rent, new firm, no image, elaboration	12%	**13.3%**	10.2%
Full-time employed, Prime credit grade A, income range $25k - $49,999k, no past due loans and no judgments, homeowner, existing firm, image , elaboration	16%	**13.5%**	19.8%
Full-time employed, Prime credit grade **A**, income range $25k - $49,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	11%	**13.7%**	22.9%
Self-employed, Prime credit grade **A**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	10%	**13.7%**	23.0%
Full-time employed, Prime credit grade **B**, income range $50k - $74,999k, past due loans and judgments, rent home, existing firm, no image , elaboration	9%	**14.5%**	42.5%
Self-employed **,** credit grade **C**, income range $25k - $49,999k, past due loans and judgments, home owner, existing firm, image, elaboration	7%	**15.6%**	25.1%
Full-time employed average credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	6%	**16.5%**	29.9%
Self-employed credit grade **D**, income range $25k - $49,999k, no delinquencies, no judgment, rent, existing firm, image, elaboration	6%	**16.9%**	33.2%
Full-time employed credit grade **E**, income range $25k - $49,999k, delinquencies, no judgment, homeowner, existing firm, image, elaboration	6%	**19.7%**	31.8%
Self-employed, credit grade **E**, income range $25k - $49,999k, delinquencies, judgment, rent, existing firm, image, elaboration	5%	**20.1%**	32.9%
Full-time employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, homeowner, new firm, image, elaboration	3%	**22.9%**	55.5%
Self-employed, high risk credit grade **HR**, income range $75k - $99,999k, past due loans and judgments, rent home, new firm, no image, elaboration	2%	**27.3%**	58.9%
Self-employed, high risk credit grade **HR**, income range $1 - $24,999, past due loans and judgments, rent home, new firm, no image, no elaboration	1%	**30.1%**	59.6%

Table 6-6: Probability of default with selection

The controls are requested loan size, offer interest rate, employment status, and income. Time, Region and industry dummies are also included in the regressions but results are not reported. Regressions are estimated using the general estimation model:

Outcome: $Pr\ (default = \alpha + \beta_1\ Owner_i + \beta_2\ Firm_i + \beta_3\ Information + \beta_4\ LoanAttributest_i + \beta_5 Industry_i + \beta_6 Macro_i + \mu$

Selection: $Pr\ (Approval) = \alpha + \beta_1 Owneri + \beta_2 Firmi + \beta_3 InformationAttributesi + \beta_4 LoanAttributes_i + \beta_5 Industryi + \beta_6 Macroi + \beta_7\ Regioni + \mathcal{E}$

Model diagnostics include the log likelihood and chi-squared statistics of the regression. Test statistics are given in parentheses. Significant coefficients are indicated with *, **, *** which stand for 0.1, 0.05, and 0.01 significance levels respectively. t-statistics in parentheses *** p<0.01, ** p<0.05, * p<0.1

	(1) Outcome (Pr default) All firms	(2) selection All firms	(3) Outcome (Pr default) Existing firms only	(4) selection Existing firms	(5) Outcome (Pr default) New firms only	(6) Granted New firms
Constant	-7.381	-0.049	-6.397	-4.921	-7.745	-0.963***
	(-0.004)	(-0.054)	(-0.004)	(-0.006)	(-0.000)	(-3.248)
Owner risk						
Credit grade (ref AA)						
A	0.409**	-0.463***	0.488**	-0.460***	0.329	-0.288***
	(2.409)	(-6.883)	(2.385)	(-5.941)	(1.573)	(-3.934)
B	0.370**	-0.712***	0.523**	-0.717***	0.262	-0.503***
	(2.027)	(-9.908)	(2.390)	(-8.611)	(1.095)	(-6.556)
C	0.411*	-1.148***	0.748***	-1.249***	0.309	-0.939***
	(1.904)	(-14.327)	(2.787)	(-13.023)	(0.971)	(-11.052)
D	0.655***	-1.460***	0.866***	-1.454***	0.486	-0.860***
	(2.927)	(-15.749)	(3.175)	(-13.356)	(1.498)	(-9.581)
E	0.696**	-1.807***	0.652*	-1.892***	0.098	-1.176***
	(2.384)	(-16.231)	(1.774)	(-14.001)	(0.235)	(-10.493)
HR	0.696**	-2.337***	0.940**	-2.396***	0.111	-1.694***
	(1.988)	(-20.443)	(2.188)	(-17.455)	(0.210)	(-15.070)
Homeowner	0.075	0.062	0.028	-0.001	0.036	0.009
	(0.703)	(1.477)	(0.223)	(-0.028)	(0.309)	(0.184)
Income Range (ref $ 0 – unable to verify)						
$1 - $24,999	-0.128	0.315***		0.219*		0.082
	(-0.576)	(3.340)		(1.907)		(0.766)
$25,000 - $49,999	-0.128	0.580***		0.497***		0.333***
	(-0.576)	(6.859)		(5.196)		(3.807)
$50,000 - $74,999	-0.230	0.581***		0.507***		0.365***
	(-1.151)	(7.586)		(5.748)		(4.488)
$75,000 - $99,999	-0.212	0.680***		0.600***		0.422***
	(-1.029)	(8.761)		(6.732)		(5.147)

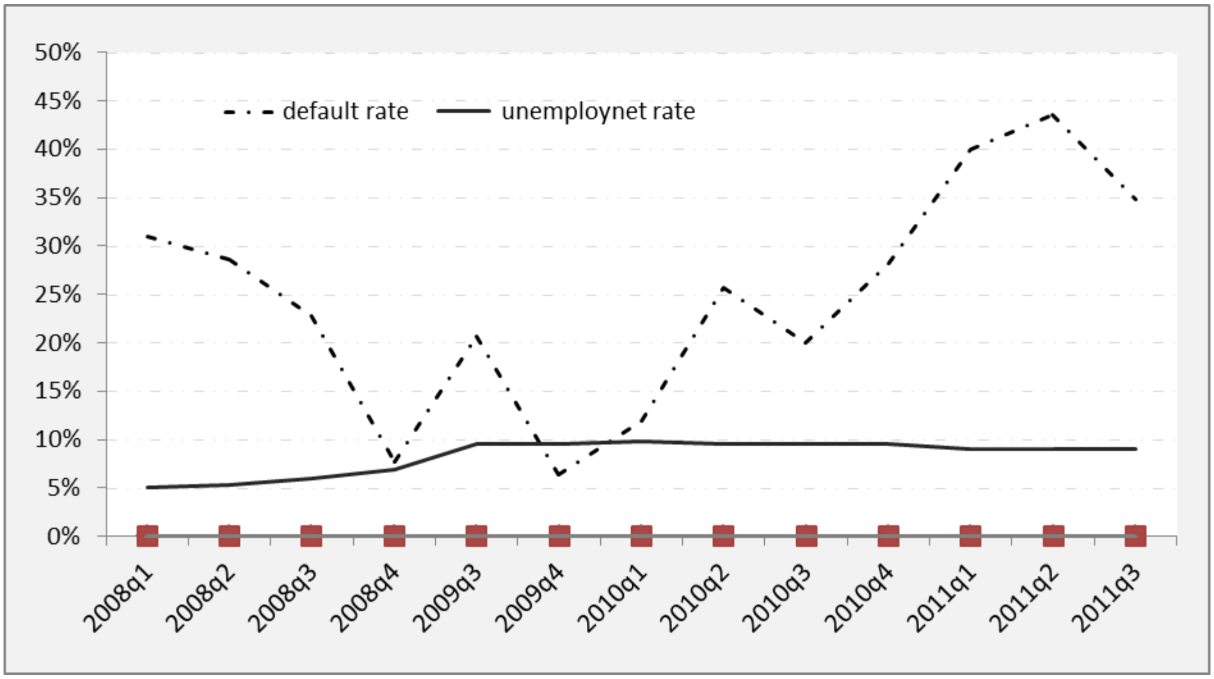
Table 6-6 continues

	(1) Outcome (Pr default) All firms	(2) selection All firms	(3) Outcome (Pr default) Existing firms only	(4) selection Existing firms	(5) Outcome (Pr default) New firms only	(6) Granted New firms
$100 000 plus	-0.184	0.642***		0.522***		0.384***
	(-0.795)	(7.346)		(5.215)		(4.219)
Firm Risk						
Existing_firm	-0.038	0.007				
	(-0.345)	(0.162)				
Loan characteristics						
Requested amount ($1,000)	0.055***	-0.164***	0.059***	-0.151***	0.027***	-0.108***
	(5.221)	(-14.005)	(4.644)	(-10.938)	(1.358)	(-8.711)
Interest rate (%)	0.051***	0.057***	0.054***	0.051***	0.037***	0.036**
	(6.225)	(4.249)	(5.462)	(3.177)	(2.695)	(2.422)
Information Attributes						
Image included	-0.289***	0.362***	-0.177	0.353***	-0.087***	0.192***
	(-2.651)	(8.091)	(-1.323)	(6.521)	(-0.640)	(4.155)
elaboration	-0.566	0.353**	-0.125	0.709***	0.257	0.594**
	(-1.205)	(2.056)	(-0.161)	(2.652)	(0.343)	(2.331)
Time dummies	Yes	Yes	Yes	Yes	Yes	yes
Industry dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	11,952	11,952	8,403	8,403	8,403	8,403
-2 log likelihood	-3157	-3157	-2264	-2264	-2419	-2419
LR test of independence	chi2(1) = 1.16 ; p=0.001					

Table 6-7: Average returns of top 5 percent lenders

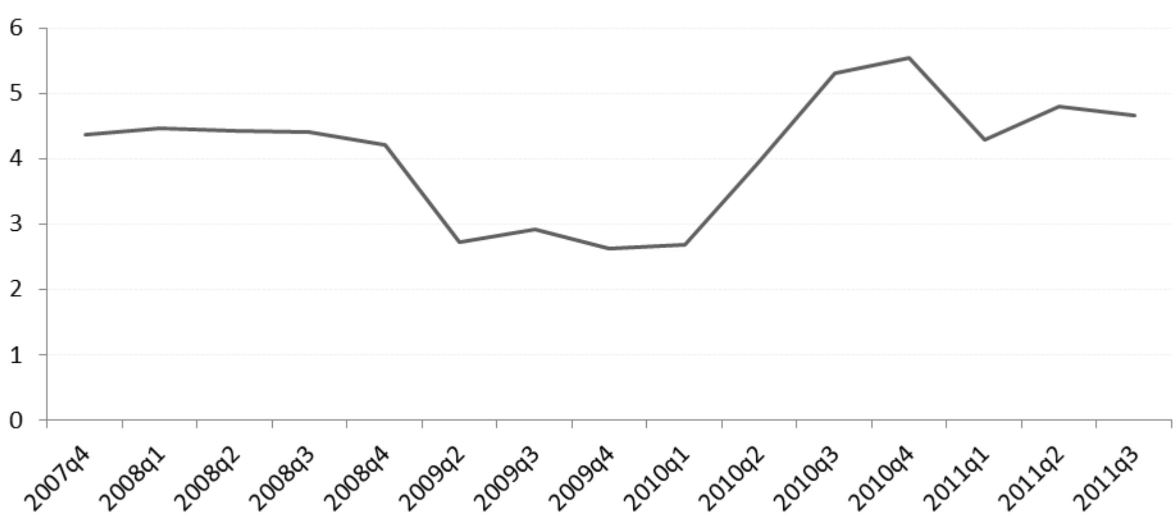
	Lender	ROI	Rate	Net Gain	Age of loans	Billed Loans	Amount Lent	Balance
1	NPX	9.3%	18.90%	$6, 408	36	23	$47, 205	$0
2	time4aloan	6.9%	16.90%	$1, 313	36	10	$13, 968	$0
3	HighonLife	6.0%	29.60%	$4, 153	36	8	$45, 288	$0
4	LoanChimp	6.0%	12.50%	$4, 076	36	10	$55, 367	$0
5	kindness-percolator5	5.9%	17.40%	$943	36	9	$12, 178	$0
6	MJARRBank	5.8%	11.50%	$1, 089	36	15	$14, 595	$0
7	Lucyqq	5.7%	9.50%	$7, 435	36	23	$124, 668	$0
8	ms-ufj	5.3%	12.20%	$1, 474	36	12	$24, 200	$0
9	moremoneymarK	5.2%	10.60%	$2, 239	36	10	$42,181	$0
10	interest-jedi0	5.0%	11.40%	$1, 308	36	26	$20,287	$0
Average results	**for top 5 percent of lenders**	**6.1%**	**15.0%**	**$3, 044.80**	**36**	**17**	**$39, 999.40**	**$0**

Figure 6-1: Average default and unemployment rates over time



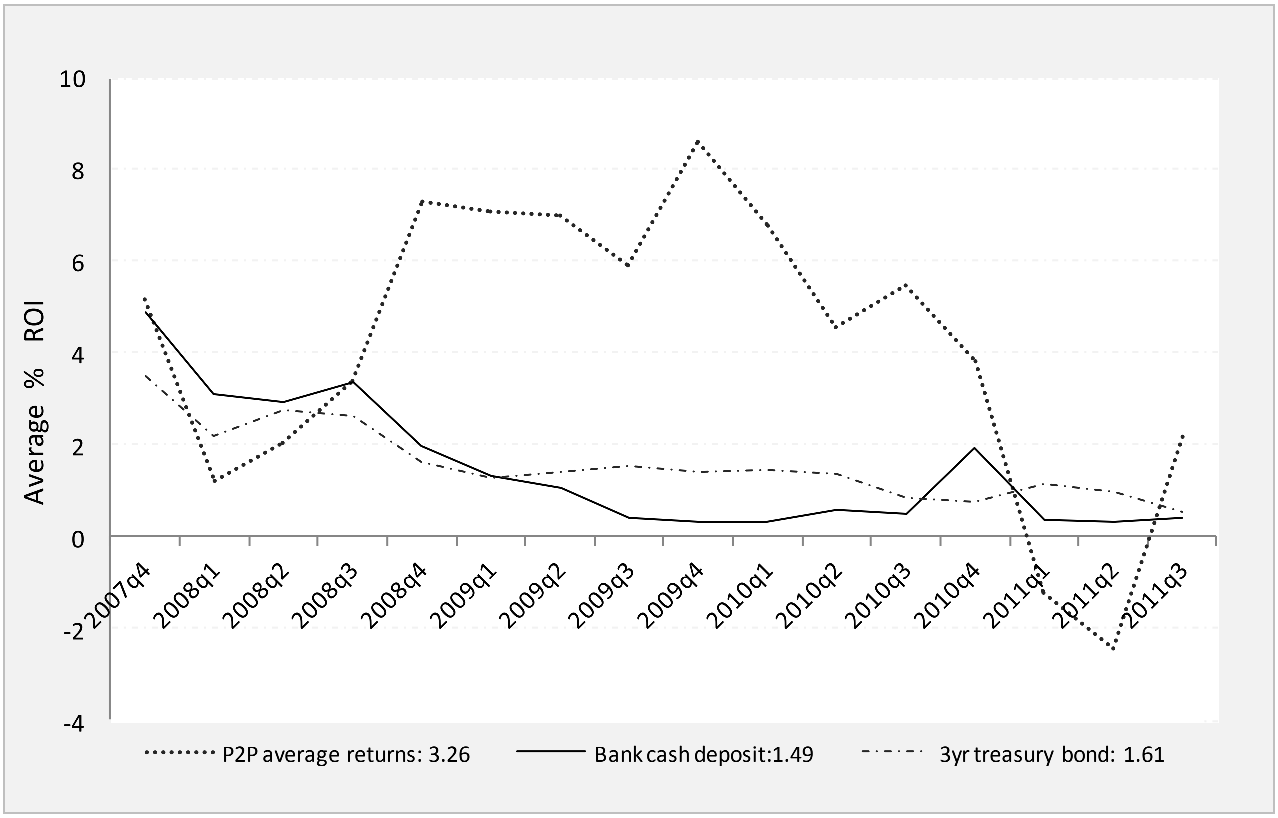
Source: Unemployment data U.S. Census Bureau 2008, 2009, 2013 www.bls.gov

Figure 6-2: Prosper credit grade proportion change over time depicting borrower risk pool



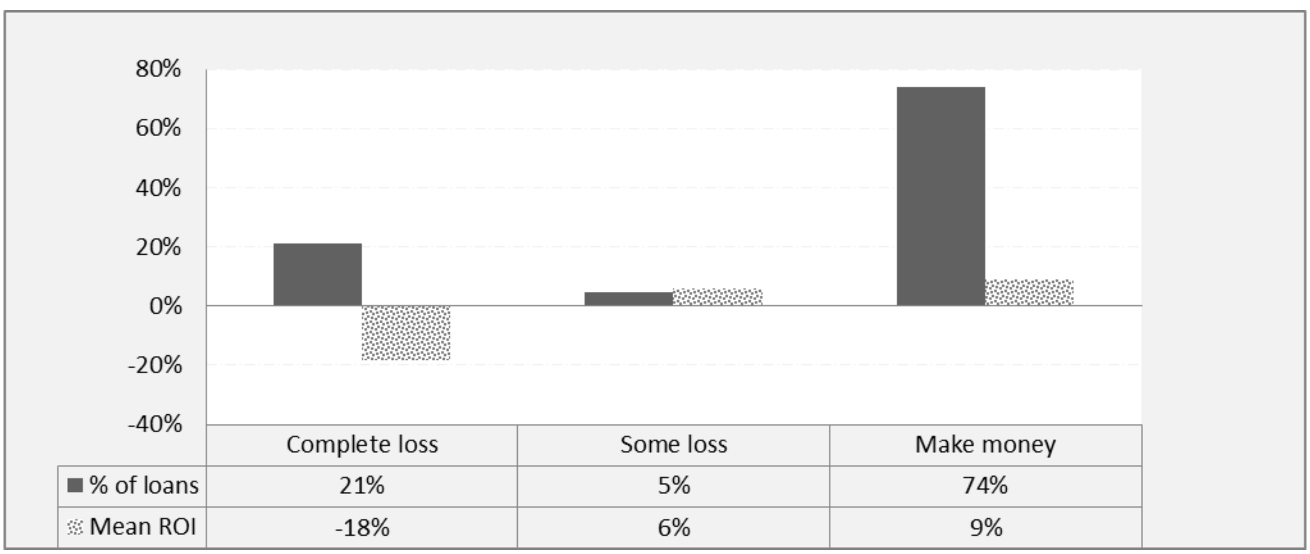
Source: www.Prosper.com 2007 -2013

Figure 6-3: P2P average returns vs. bank cash deposit vs. treasury bonds



Source: Cash deposit data 2008-2013 www.bankrate.com ; Treasury bonds data 2008-2013 www.treasury.gov

Figure 6-4 Percent distribution of returns for all funded loans



	Complete loss	Some loss	Make money
% of loans	21%	5%	74%
Mean ROI	-18%	6%	9%

Figure 6-5: Distribution of returns by loan size

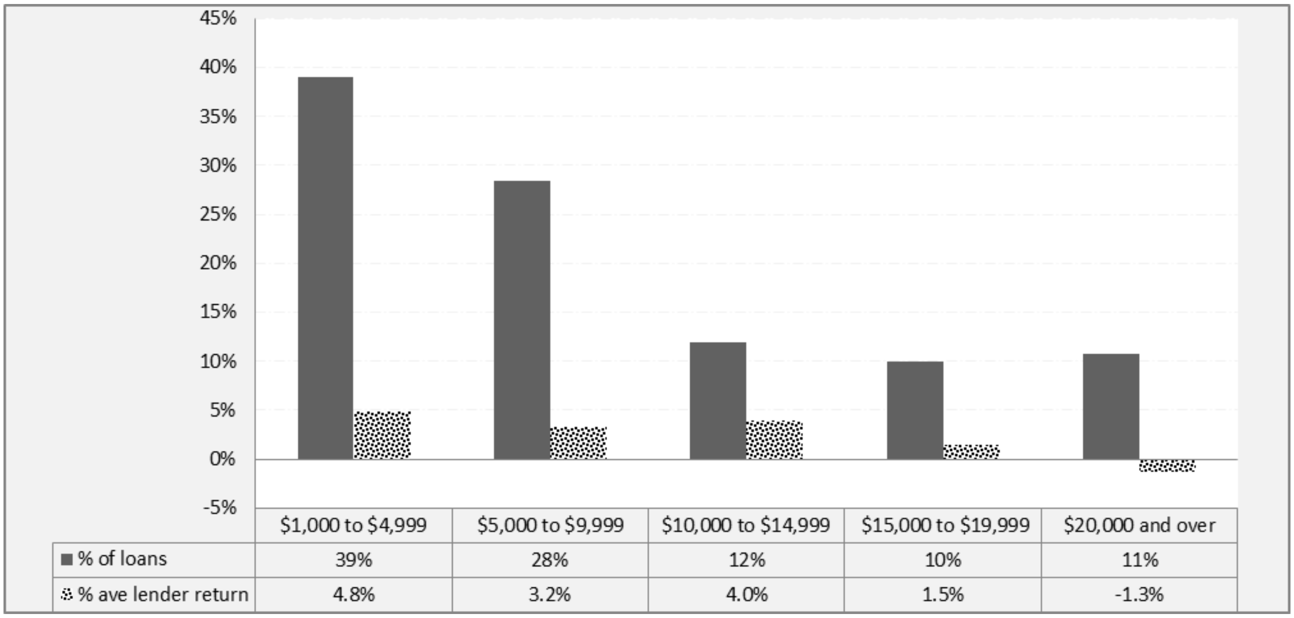


Figure 6-6: Distribution of returns by credit rating

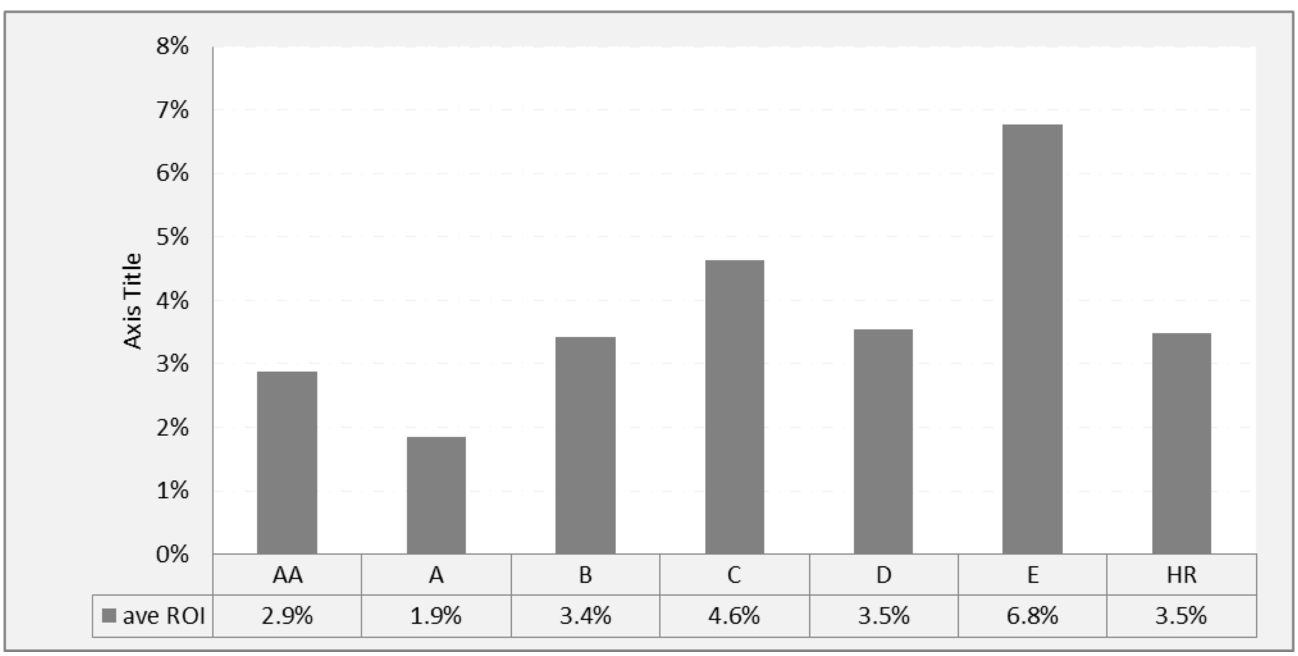
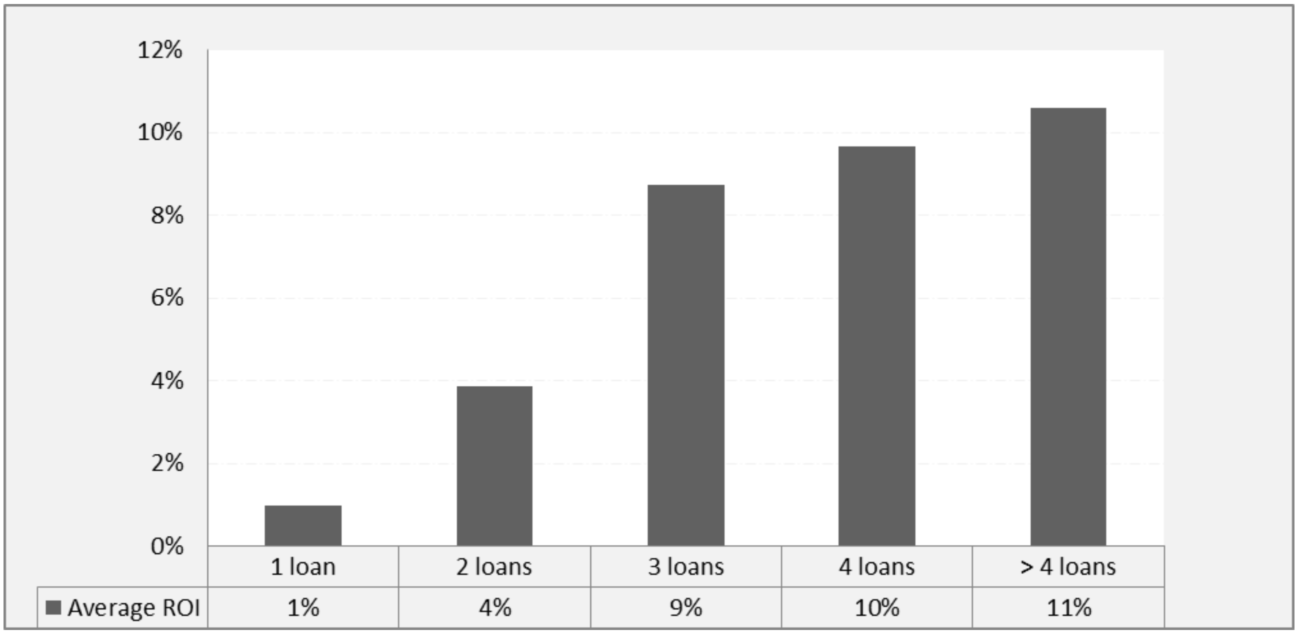
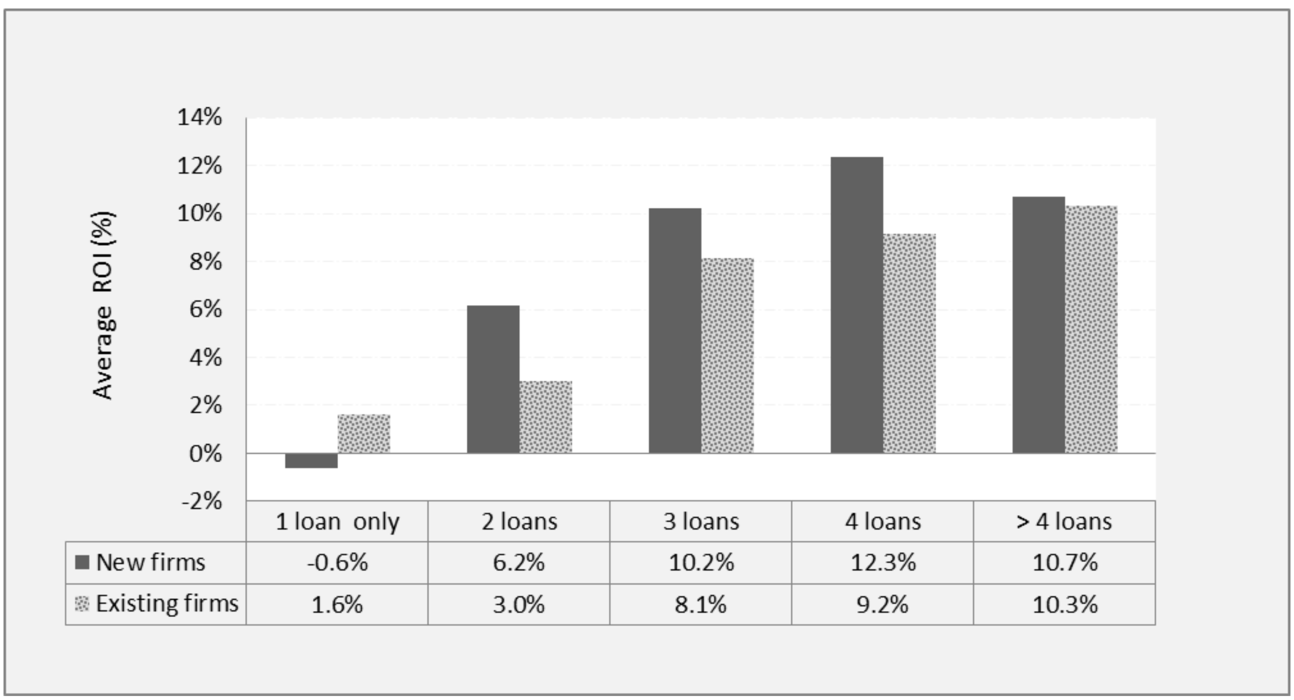


Figure 6-7: Are there repeat winners?



	1 loan	2 loans	3 loans	4 loans	> 4 loans
Average ROI	1%	4%	9%	10%	11%

Figure 6-8: Distribution of returns by Firm Age



	1 loan only	2 loans	3 loans	4 loans	> 4 loans
New firms	-0.6%	6.2%	10.2%	12.3%	10.7%
Existing firms	1.6%	3.0%	8.1%	9.2%	10.3%

Chapter 7 Conclusions, contributions and limitations

7.0 Introduction

Differences in the severity of information problems may help explain why certain types of small businesses obtain external finance from traditional lending institutions, whereas other types tend to rely more on informal investors. Of cause information problems may also keep firms from obtaining external funding at all. In this dissertation, we set out to examine empirically the impact of P2P lending in the financing of small business ventures. We follow entrepreneurial finance literature in raising three key questions concerned with credit extension; cost of credit and loan default. We gathered and created a unique data set taken from Prosper.com, one of the dominating P2P lending websites; which formed the basis of our analysis.

Overall, one of the key findings of the study shows that P2P lending depicts a new small business venture loan market, where previously underserved early stage entrepreneurs and those looking for small amounts are able to access unsecured credit through the relaxation of collateral. Although collateral is not required, we show that the supply of loans tends to flow to the least risky entrepreneurs; those who are homeowners, with high credit ratings. In our findings, we also show that firm level characteristics (including the age of the firm), have little impact on loan supply while reducing information asymmetries through giving volunteering information improves access to loans. In general, our findings are both interesting and important as they suggest that P2P lending is a low risk form of debt finance. In this sense, lenders act like traditional debt financiers. However, the way in which they appraise funding opportunities characterise typical decision making of equity investors such as Business Angels and VC, who tend to focus more on people, rather than the business itself.

In terms of the cost of credit, we observe that with average lending of between 18 percent and 20 percent, P2P lending is a very expensive form of debt finance. Banks typically refuse to extend credit given such high interest rates as this tends to alter the borrower pool such that only the riskiest of borrowers have projects that generate returns that are high enough to be able to re-pay these interest rates. In effect, the bank supply curve is backward bending above 10 percent on conventional terms of lending. Consequently, if we were to characterise P2P lending we would effectively conclude that it is typically a high cost finance with required returns expected to be likely in the levels of Business Angels and VC equity investments.

So, if within the P2P lending market it is the relatively low risk borrowers who get the loans, then a natural question to address is why couldn't these low risk borrowers access conventional bank loans? Moreover, why were they willing to offer high interest rates? One probable explanation is the condition of the credit market. Our analysis is based on data covering the period 2008 - 2013, arguably characterised as the height of the recession. According to SBA (2012) small business lending from banks constricted during this period. Suddenly, businesses who were perfectly good risk able to access credit from banks during buoyant economic times, were suddenly unable to tap into bank lending because banks tightened their lending standards to small firms. Subsequently, borrowers, keen to have access to capital, were hoping to attract possible lenders with their good credit record and high interest rates; just so they can keep their businesses afloat. Consequently, we find that P2P lending platforms provided an alternative form of venture finance in times of recession, when traditional financiers were unavailable.

In terms of lender return and default, we find that the expected return to lenders is 3.26 percent, which is above the opportunity cost of capital in the US. Therefore, P2P lending is profitable from the investor point of view, albeit in a narrow sense. In general, the results suggest that average lenders on P2P platforms are amateurs, who actually have a higher risk tolerance. For these lenders, the risk of losing a small proportion (as little as $25) per investment in the overall portfolio of loans is offset by the potential gain from high interest rates charged for loans. However, default is found to be related to risk as predicted by conventional theory, in both formal credit rating and also additional information and track records. Moreover, a key finding is that default seems to also be associated with the cost of credit, which presents an interesting dilemma for lenders. Consequently, at a first glance, we inferred from our finding that lenders would have to consider a potential trade off between lowering the interest rates they charge borrowers in order to circumvent default. However, it is also plausible that some lenders extend credit to firms in this context for fun - as some form of 'gambling'. Hence it would not be necessary to consider a potential trade off between interest rates and default. In this case, lenders would simply accept loan requests from borrowers offering high interest rates with the view that knowing that if the borrowers pay back the loan, they win big. However, if some of the loans in the portfolio result in default, the loss is not too big.

Furthermore, it is also reasonable to consider that some of the lenders choose to extend credit in this context driven by philanthropic reasons. This is supported by evidence we found earlier that lenders in this market are forgiving – there is a subset of low risk borrowers with previous delinquencies, who have managed to rebuild their credit history, extended funding in this context. Therefore, lenders driven by philanthropy might not necessarily have wealth maximisation as his main goal for funding these firms.

Interestingly, our results show that return from the top 5 percent of lenders average at 6.1 percent per annum. Given the fact that P2P lending is generally a young market, and the fact that majority of lenders attracted to P2P lending are relatively uninformed amateurs in making investment decisions, the results suggest that if the amateur lenders do indeed learn, it then becomes plausible that in time the returns in this market may generally converge to be better (and gravitate towards the 6.1 percent achieved by top 5 percent). However, if the P2P lending platforms continue to attract a pool of amateur lenders, the average returns of 3.26 percent may render the market somewhat unsustainable in the long run. Given all that has been presented so far, P2P lending may come across as an inefficient use of resources. For example, if we were to follow the line of argument that lenders may choose to engage in P2P lending for idiosyncratic reasons such as fun or gambling etc and philanthropic reasons, it then becomes reasonable to infer that perhaps P2P lending may possibly result as an inefficient use of resources. However, the fact that P2P lending is profitable for an average, risk loving, uninformed lender, may suggest that this instrument may somewhat be an alternative form of financial instrument (albeit with some form of inefficient tendencies).

7.1 Key contributions

So how do our results inform theory? To recap, Stiglitz and Weiss put forward four key parameters underpinning their theory of information asymmetry problems, namely: conducting due diligence, collateral, refraining from high interest rates to avoid moral hazard and adverse selection issues, and the inferred face to face borrower-lender interactions. From our results, we find that some weights of these parameter values are likely to change in Stiglitz-Weiss model.

Firstly, the general insight we get from our research is that due diligence, although still an important factor, in the P2P lending context it is conducted by the ‘crowd’. This new feature, unique to P2P lending was not taken into consideration in the Stiglitz and Weiss framework –

where credit risk appraisal was done by relatively one person. Consequently we introduce collecting intelligence as a means of eradicating information asymmetry issues. Furthermore, our results shift focus from 1-to-1 physical interactions between borrowers and lenders inferred in Stiglitz-Weiss theory and highlights the 1-to many borrower lender typology over the internet. Effectively rendering physical contact, which was previously seen an important aspect of reducing information asymmetry issues in theory, relatively less important.

Furthermore, In terms of due diligence, the crowd also introduces another distinctive change based in the notion that lenders in this context may have philanthropy ambitions which they may consider when appraising credit risk. For example, a lender taking into consideration philanthropic ambitions may look for different credit risk when compared to a lender whose sole ambition is to maximise returns. This may effectively alter access to credit, the cost of credit and subsequently the P2P borrower pool.

Second, we see from our results that collateral is being taken out; which was such an important determinant in reducing both adverse selection and moral hazard issues in Stiglitz-Weiss theory. In the P2P lending it is unimportant.

Third, Similar to theory, a key finding in our analysis is that cost of capital is associated with default. However, this finding represent some sort of interesting dilemma for lenders in the P2P lending context, given that lenders do not seem to be easily put off by the high interest rates offered by prospective borrowers and the relative default risk. Our results suggest that these lenders are more risk loving, with a different utility to that of the traditional banks (typically concave) discussed in Stiglitz and Weiss theory. Unlike the 1-to-1 lender borrower typology of traditional banks, where high interest rates may result in huge losses; in the P2P lending market the many-to-1 lender borrower typology result in the idea of fun being introduced; where lenders may be extending credit because they like it especially given the fact that if the investment fails, the losses may not be that big.

Fourth, the general insight we get from our study is that borrower reputation, stipulated by credit grades, is the single most important determinant of credit allocation, the cost of capital and loan default predictions. The significant of using the credit grade helps to reduce the problem of selecting credit worthy risk and reducing moral hazard. The cost of defaulting will result in poorer scores – which are quantifiable to 24 percent in reduction of probability of funding and increase of 80 basis points interest rates for an otherwise average risk borrower. With the advancement of internet, reputations which were previously curtailed within the 1-to-1 lending typology from banks, where if a borrower defaults on credit in one region or country for example, would not have an effect if the borrower were to move to another country has since have completely changed. But with internet age loan default may quickly go viral. The consequence of 1-to- many borrower-lender interactions relative to reputation over the internet and the ease with which default may go viral makes reputation to be a very important aspect within the P2P lending context.

In sum, we contend that the lessons we leant from our results about asymmetric information issues are quiet different to those developed in Stiglitz-Weiss model. Reputation is very important in this market (this was not highlighted by Stiglitz). We learn that physical contact and collateral becomes less important in reducing information asymmetries. We also learn about the importance of three key new features collective intelligence of the crowd, an aspect of philanthropy present when appraising credit risk and the general element of fun which may drive lenders when choosing to allocate and price credit. Their model will have to be updated to take into consideration the facts raised above in their theory of information asymmetry in order to reflect our finding.

Our results also have key practical implications. The fact that P2P lending relaxes collateral requirements and the fact that reputation is the single most important variant in P2P lending renders our results generalizable to the microfinance institutions, especially those in developing countries. The relaxation of collateral makes P2P a viable alternative to small business funding in developing countries, especially given the fact that it is in these regions where the wealth distributions tend to be skewed unfavourably. The importance of reputation simply highlights infrastructure that already exists, used by microfinance in developing countries. Hence this renders our results somewhat generalizable. The fact that the internet is an underlying layer in operationalizing P2P lending in developing countries, perhaps there is scope of using mobile smart phones instead of traditional computers – given that this technology is already entranced in developing countries.

7.2 Limitations

Like most studies that use secondary data, our data were not specifically collected for small business study; hence we have some limitations. First, ideally, any analysis of small business funding ought to include those applicants who chose not to apply for credit - even though they could potentially qualify for credit (Hanley and Germa, 2006).

Second, the data do not contain borrower demographic variables and human capital variables. Previous studies (see for example Coleman, 2000; Cressy, 1996; Burke *et al,* 2000) have found these to be important in influencing credit access and the cost of credit. Our study is therefore not immune to omitted variable bias.

Third, majority of our data (in terms of loan volumes) comes from the period 2008; the nearest crisis and post-crisis period may not be very representative, but it stimulated an innovative ways of interrelation between lenders and borrowers. Hence, as part of future research, it would be interesting to see whether factors driving the likelihood of attaining small business funds in this market vary in more cheerful – post crisis market conditions (Cowling et al, 2012).

Overall, in terms of our finding, we interpret two of our key findings with caution. First, is the interpretation of the variable *Bid_count*; which we use to measure the level of information asymmetries reduced by the intelligence of the crowd. In this study, we simply use the count (number) of lenders that extend credit as an indicator of some form of due diligence coming through. It is possible that other effects may actually be at play. For, instance, herding behavior may be a possible explanation; where lenders gravitate to loan requests based on other features that we are unaware of and hence did not investigate here. For instance, there are blogs that lenders use to share information and communicate about the loan requests, which were not available to the researches due to regulatory issues; hence they were unavailable in our data. Since, P2P context continues to evolve, some of the data has since been made public to anyone who registers and joins the site as a lender. Therefore one area of future research would be to look closer at the collective intelligence of the crowd and ascertain more precisely what forces are at play use to overcome problems lending under asymmetric information.

Second, the same caution goes for the elaboration variable. Due to the sheer volume of the number of cases (N =12, 526); we have created this variable as a binary indicator – showing whether the small business borrower tells a story or not. A better measure would have been to disentangle and look at the actual contents of the story. Moreover, borrowers also had private conversations with lenders in the form of Questions/Answers; where lenders could ask borrowers any questions based on their loan request. Similarly, these were not made public due to regulatory constraints. Hence they are unavailable for our data. However, some of these restrictions have since been lifted such that any lender, that registers on the platform they can have access to full disclosure of any conversation that potential borrowers may have had with prospective lenders. This would make an interesting extension of the research. Finally, since this market is new and continues to grow, there will be more opportunity to track the data and see how the return evolves over time.

Research dissemination

In terms of disseminating our research to a wider academic community, we have developed two empirical papers from the three chapters, which have been disseminated through academic conferences as follows:

- Paper titled: "New technology same old story: factors driving credit allocation for small business loan on commercial Peer-to-Peer lending websites" accepted at the 6th Annual conference academy of innovation and entrepreneurship (AIE, 2013), Oxford University, UK

- Presenter at policy conference: **"**New technology same old story: factors driving credit allocation for small business loan on commercial Peer-to-Peer lending websites" 11th Annual International Conference on Finance, Athens, Greece (2013)

- Presenter at policy conference, "What's in store for tomorrow's SME finance: the case of Crowd Funding", *Strategies to Overcome Poverty and Inequality: Towards Carnegie 3* , University of Cape Town, South Africa (2012)

- Presenter paper titled: Research methods in entrepreneurship: "what determines the endowment of quality entrepreneurs in an economy",19th European Doctoral Programme Association in Management and Business Administration Seminar, Soreze, France (2010)

Table 7-1: Summary of findings

This table shows the developed hypotheses, predicted relationships and what we found as determinants of: credit allocation, the cost of credit and loan default for P2P small business loans.

Hypothesis		Predicted relationship	Found relationship
Factors driving credit approval			
H_{1a}	small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are more likely to be extended credit by prospective lenders	+	+
H_{2a}	small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are more likely to be extended credit by prospective lenders	+	+
H_{3a}	small business borrowers who have successfully paid back a previous loan are more likely to be extended credit	+	+
H_{4a}	existing firms are more likely to be funded (relative to new business start-ups) in the P2Plending context	+	Variable insignificant
H_5	the likelihood of a partially funded loan receiving additional credit will increase with the total number of already extended offers	+	+
H_{6a}	small business borrowers who use text elaborations are more likely to be funded in the P2P lending context	+	Variable insignificant
H_{7a}	small business borrowers who have previous failures are less likely to access funds in the P2P lending context	-	-
H_{8a}	small business borrowers who post pictures are likely to get funded by lenders in the P2P lending context	+	+
Factors driving cost of credit			
H_{1b}	small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are more likely to pay lower interest rates	-	Variable insignificant
H_{2b}	small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are more likely to pay lower interest rates	-	-
H_{3b}	small business borrowers who have successfully paid back a previous loan are more likely to pay lower interest rates	-	Variable insignificant
H_{4b}	existing firms are more likely to pay lower interest rates (relative to new business start-ups) in the P2Plending context	-	Variable insignificant
H_{6b}	small business borrowers who use text elaborations are more likely to be pay cheaper interest rates in the P2P lending context	-	Variable insignificant
H_{7b}	small business borrowers who have previous failures are more likely to pay higher interest rates	+	Variable insignificant
H_{8b}	small business owners who post pictures are likely to pay lower interest rates in the P2P lending context	-	-
Factors driving default activity			
H_{1c}	small business borrowers, who own their homes, demonstrate better credit risk (relative to those that rent) and therefore are less likely to default	-	Variable insignificant
H_{2c}	small business borrowers, with high credit ratings, demonstrate better credit risk and therefore are less likely to default	-	-
H_{3c}	small business borrowers who have successfully paid back a previous loan are less likely to default	-	-
H_{4c}	existing firms are less likely to default (relative to new business start-ups) in the P2P the lending context	-	Variable insignificant
H_{6c}	small business borrowers who use text elaborations are more (less) likely to be default	+/-	Variable insignificant
H_{7c}	small business borrowers who have previous failures are more likely to default	+	Variable insignificant
H_{8c}	small business owners who post pictures are more (less) likely to default	+/-	-

Chapter 8 References

Agarwal, S; Chomsisengphet, S and Liu C (2007), "Determinants of small business default", In: G Christodoulakis, S Satchell (Eds), the analytics of risk model validation, *Palgrave-Macmillan Publishing*, pp. 1–12

Agrawal, A; Catalini, C and Goldfarb, A (2011),"The Geography of Crowdfunding", NBER *Working Paper*, No. 16820

Ang, J (1991), "Small Business Uniqueness and the Theory of Financial Management", *The Journal of Small Business Finance*, vol.1, pp.1 - 13

Ang, J (1992), "On the Theory of Finance for Privately Held Firms", *The Journal of Small Business Finance*, vol.1, pp.185 - 203

Altman, E and Sabato, G (2007), "Modeling credit risk for SMEs ", Abacus, vol.43, no.2, pp. 332 - 357Arrow, K (1973), "The theory of Discrimination", in Aschenfelter, O and Rees, A. (ed.), Discrimination in Labor Markets: *Princeton University Press*

Amatucci, F. M. and J. E. Sohl (2007), "Business angels: investment processes, outcomes and current trends, in A. Zacharakis and S. Spinelli Jr. (eds.), Entrepreneurship: The Engine of Growth, Volume 2, Westport CT, Praager, pp. 87–107

Ashta, A. and Assadi, D (2010), "Online or Offline: The Rise of Peer-to-Peer Lending in Microfinance", *Journal of Electronic Commerce in Organizations*, vol. 8, no. 3, pp. 26-37

Athavale, M and Edmister, R (2004), "The pricing of sequential bank loans", *Financial Review*, vol.39, pp.231-253

Avery, R; Bostic, R and Samolyk, K (1998), "The role of personal wealth in small business financing", *Journal of Banking and Finance*, vol.22, pp.1019 -1061

Barasinska, N (2010) "Does Gender Affect the Risk Propensity of Retail Investors? Evidence from the Largest German Market for Peer-to-Peer Lending", Thesis

Barasinska, N (2011), "Does Gender Affect Investors' Appetite for Risk? Evidence from Peer-to-Peer Lending" DIW Berlin *Discussion Paper* No. 1125 Available at SSRN: http://ssrn.com/abstract=1858719

Baum, J and Silverman, B (2004), "Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology start-ups", *Journal of Business Venturing*, no.8, pp. 25-48

Bates, T (1991), "Commercial bank financing of white and black owned small business start-ups", *Quarterly Review of Economics and Business*, vol. 31 no.1, pp. 64-80

Benjamin, G and Margulis, J (2000), " Angel financing: how to find and invest in private equity", New York, John Wiley and Sons Publishers

Berger, A and Udell, G (1990), "Collateral, loan quality and bank risk", *Journal of Monetary Economics*, vol. 25, pp. 21 – 42

Berger, A and Udell, G (1992), "Some evidence on the empirical significance of credit rationing", *Journal of Political Economy*, vol. 100, no. 10, pp. 1047-1077

Berger, A and Udell, G (1995),"Relationship lending and lines of credit in small firm finance", *Journal of Business*, 68, pp.51-381

Berger, A and Udell, G (1998), "The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle", *Journal of Banking and Finance*, vol. 22, no. 6-8, pp. 613-673

Berger, A; Frame, S and Miller, N (2005), "Credit scoring and the availability, price, and risk of small business credit", *Journal of Money, Credit and Banking*, vol.37, no2, pp. 191–222

Berger, A; Cowan, A and Frame, S (2011), "The surprising use of credit scoring in small business lending by community banks and the attendant effects of credit availability, risk and profitability", *Journal of Financial Services Research*, vol. 39, pp. 1 - 17

Berger, A and Frame, S (2007), "Small Business Credit Scoring and Credit Availability", *Journal of Small Business Management*, vol. 47, pp. 5-22

Berger, S and Gleisner, F, (2009), "Emergence of financial intermediaries in electronic markets: The case of online P2P lending", *Business Research*, vol. 2, no. 1, pp. 39-65

Behr, P; Guttler, A and Plattner, D (2004), "Credit scoring and Relationship lending: the case of Germany SMEs, mimeo

Besanko D and Thakor, A (1987), "Collateral and rationing: sorting equilibria in monopolistic and competitive credit markets ", *International Economic Review*, vol.28, no.3, pp.671-689

Bester, H (1985), "Screening vs. rationing in credit markets with imperfect information ", *The American Economic Review*, vol.75, no.4, pp. 850-855

Bester, H (1987), "The role of collateral in credit markets with imperfect information", *The European Economic Review*, vol.31, pp. 887-899

Binks, M and Ennew, C (1998), "Private Businesses and their Banks", Forum of Private Business Bank Survey, 1998, University of Nottingham Business School

Blanchflower, D and Oswald, A (1998), "What Makes an Entrepreneur?" *Journal of Labor Economics*, vol. 16, no. 1, pp. 26

Boot, A and Thakor, A (1994), "Moral hazard and secured lending in an infinitely repeated credit market game", *International Economics Review*, vol. 35, no. 11, pp. 899 - 920

Boot, A; Thakor, A and Udell, G (1991), "Secured Lending and Default Risk: Equilibrium Analysis, Policy Implications and Empirical Results", *Economic Journal*, vol. 101, no. 406, pp. 458-472

Brick, I and Palia, D (2005), " Evidence of jointness in the terms of relationship lending", Rutgers University, *Working Paper*

Burke, A and Hanley, A (2006), "Bank interest margins and business start-up collateral: testing for convexity", *Scottish Journal of Political Econom*y, vol. 53, no. 3, pp. 319-34

Burke, A and Hanley, A, (2003), "How do firms pick safer ventures? A theory relating the importance of risk aversion and collateral to interest rates and credit rationing", *Journal of Entrepreneurial Finance and Business Ventures*, vol. 8, 2, pp. 13 - 24

Burke, A; FitzRoy, F and Nolan, M (2000), "When Less is More: Distinguishing between Entrepreneurial Choice and Performance", *Oxford Bulletin of Economic and Statistics*, vol. 62, no. 5, pp. 565-587

Burtch, G; Ghose, A and Wattal, S (2011)"An empirical Examination of the Antecedents and consequences of investment patterns in Crowd- Funded markets", SSRN Electric Journal, doi:10.2139/ssrn.1928168

Board of Governors of Federal Reserve System (2012), "Availability of credit to small businesses", http://www.federalreserve.gov/publications/other-reports/files/sbfreport2012.pdf, report last accessed 22 April 2013

Bygrave, W. and Quill, M. (2007), “Global Entrepreneurship Monitor 2006 Financing Report”, Wellesley-London: Babson College and London Business School, Global Entrepreneurship Research Association

Bygrave, W. D. and Reynolds, P. D. (2004) "Who finances start-ups in the USA? A comprehensive study of informal investors, 1999-2003,"In Zahra, S. A. et al. (eds.), Frontiers of Entrepreneurship Research 2004, pp. 37-47, *Wellesley, MA: Babson College*

Bygrave, W; Hay, M and Reynolds, P (2003), "Executive forum: a study of informal investing in 29 nations composing the Global Entrepreneurship Monitor", *Venture Capital*, vol. 5, no. 2, pp. 101

Cassar, G (2004), "The financing of business start-ups*", Journal of Business Venturing*, vol. 19, pp. 261-283

Chatterjee, S; Corbae, D; Nakajima, M and Rio-Russ, J (2007),"A quantitative theory of unsecured consumer credit with risk of default", Econometrica, vol. 75, pp.1525 - 1589

Chemin, M and De Laat, J (2009),"Can Warm Glow Alleviate Credit Market Failure? Evidence from Online Peer-to-Peer Lenders", *CIRPEE Working Paper 09-34*. Available at SSRN: http://ssrn.com/abstract=1461438

Cope, J; Cave, F and Eccles, S (2004), "Attitude of venture capital investors towards entrepreneurs with previous business failure", *Venture Capital*, vol.6, pp 147- 172

Cowling, M (1997), "Small Business Loan Requests: who gets the cash?", SME Centre *Working Paper* No.52, Warwick Business School

Cowling, M (1999), "The determination of bank-small business loan premia in the UK" *Applied Economics letters*, vol.65, pp251-253

Cowling, M, Liu, W and Ledger, A (2012) Small business financing in the UK before and during the current financial crisis, *International Small Business Journal*, Online first, 4 July

Cowling, M and Mitchell, P (2003), "Is the Small Firms Loan Guarantee Scheme Hazardous for Banks or Helpful to Small Business?” *Small Business Economics*, vol. 21, no. 1, pp 63–71

Cole, R (1998), "The importance of relationships to the availability of credit", *Journal of Banking and Finance*, vol. 22, pp. 959-997

Coleman, S (2000),"Access to capital and terms of credit: A comparison of men and women owned small businesses", *Journal of Small Business Management*, vol.20, no.3, pp.37-52

Cosci, S and Meliciani, V (2002), " Multiple banking relationships: Evidence from the Italian experience", *Manchester School*, vol. 70, pp 37-54

Cressy, R. (1996), "Commitment lending under asymmetric information: theory and tests on UK start-up data", *Small Business Economics*, vol.8, no. 5, pp. 397 - 408

Cressy, R. (1996), "Are business start-ups debt-rationed", Economic Journal, vol. 106, no. 438, pp. 1253-1270

Cressy, R (2002), "Introduction: Funding Gaps", Economic Journal, vol. 112, no. 477

Cressy, R. and Toivanen, O. (2001), "Is there adverse selection in the credit market?", *Venture Capital*, vol. 3, no. 3, pp. 215-238

Damodaran, A (2010), "The Dark Side of Valuation: Valuing Young, Distressed, and Complex Business", 2nd Edition, *Pearson Education*

Darlington, R.D (1990), Regression and Linear Models, *New York: Wiley*

Deakins, D; North, D; Baldock, R, and Whittam, G (2008), "SMEs access to finance: is there still a debt gap?", paper presented to the Institute for Small Business and Entrepreneurship Conference, Belfast, November

Degryse, H and Van Cayseele, P (2000), " Relationship lending within a bank-based system: evidence from European Small Business data", *Journal of financial intermediation*, vol.9, pp. 90-109

De Meza, D and Southey, C (1996), " The borrower's curse: optimism, finance and entrepreneurship", *Economic Journal*, vol. 106, no. 435, pp. 375 - 386

Dennis, S; Nandy, D and Sharpe, I (2000), " The determinants of contract terms in bank revolving credit agreement", *Journal of Financial and Quantitative Analysis*, vol.35, pp. 87-110

DeYoung, R; Glennon, R and Nigro, P (2008), "Borrower-lender distance, credit scoring, and the performance of small business loans", *Journal of Financial Intermediation*, vol.17, pp. 113–143

Diamond, D (1989), "Reputation acquisition in Debt Markets", *Journal of Political Economy*, vol. 97, no. 4, pp. 828-862

Duarte, J., Siegel, S and Young, L (2010), "Trust and Credit", AFA 2010 Atlanta Meetings, Paper available at SSRN: http://ssrn.com/abstract=1343275

Dunne, T; Roberts, M. and Samuelson, L (1989), " The growth and failure of US manufacturing plants", *Working Paper 1875*, Pennsylvania State, Department of Economics

Dyrberg-Rommer, A (2005), "A comparative analysis of determinants of financial distress on French, Italian and Spanish firms", *Working paper series 26*, Denmark's National bank

Evans, D. S. and Jovanovic, B. (1989), "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints", *Journal of Political Economy*, vol. 97, no. 4, pp. 808

Evans, D. (1987), "The relationship between firm size and age", *The Journal of Industrial Economics*, vol. 35

Everett, C. R (2010), "Membership, Relationship Banking & Loan Default Risk: The Case of Online Social Lending", Available at SSRN: http://ssrn.com/abstract=1114428

Fidrmuc, J and Hainz P (2009), "Default rates in the loan market for SMEs: Evidence from Slovakia, IFO *Working Paper*, No 72, http://ideas.repec.org/p/wdi/papers/2006-854.html

Freedman, S and Jin, G. Z (2008), "Do Social Networks Solve Information Problems for Peer-to-Peer Lending? Evidence from Prosper.com" *NET Institute Working Paper No. 08-43*. Available at SSRN: http://ssrn.com/abstract=1304138

Glennon, D and Nigro, P (2008), "Measuring the Default Risk of Small Business Loans: A Survival Analysis Approach", *Journal of Money, Credit and Banking*, vol.3, pp.923-947

Glennon, D. and Nigro, P (2005), "An Analysis of SBA Loan Defaults by Maturity Structure", *Journal of Financial Services Research*, vol. 28: pp. 77-111

Good, W and Graves, R (1993), "Small business support programs: the views of failed versus surviving firms", *Journal of Small Business and Entrepreneurship*, vol. 10, pp. 66-76

Greene, W. H (1997), Econometric Analysis (3rd Ed.), *New York: Prentice Hall*

Hall, J (1993), "Venture Capitalists' decision criteria in new venture evaluation", *Journal of business venturing*, no.8, pp. 25-42

Han, L and Greene, F (2007), "The determinants of online loan applications for small businesses", *Journal of Small Business and Enterprise Development*, vol. 14, no.3, pp. 478 - 486

Hanley, A and Crook, J (2005), "The higher cost of follow up loans ", *Small Business Economics*, vol.24, pp 29 – 38

Hanley, A and Girma, S (2006), "New ventures and their credit terms", *Small Business Economics,* vol.26, pp 351-364

Harhoff, D and Körting, T (1998), "Lending relationships in Germany: Empirical results from survey data", *Journal of Banking and Finance*, vol.22, pp 1317 -1354

Hartley, S (2010), "Crowd–Sourced microfinance and cooperation in group lending", Working Paper, Social Science Research Network, available as papers.ssrn.com/sol3/papers.cfm?abstract_id=1572182

Heckman, J (1978), "Dummy endogenous variables in a simultaneous equation system", *Econometrica*, vol. 46, no4, pp. 931–959

Heminway, J, and Hoffman, S (2010), "Proceed at your peril: crowdfunding and the securities act of 1933", *Tennessee Law Review*, vol, 78, pp.879–972

Herzenstein, M., Dholakia, U. M., and Andrews, R. L (2011), "Strategic herding behavior in peer-to-peer loan auctions", *Journal of Interactive Marketing*, 25(1), 27-36

Herzenstein, M., Sonenshein, S and Dholakia, U. M (2011), "Tell Me a Good Story & I May Lend You My Money: The Role of Narratives in Peer-to-Peer Lending Decisions", Available at SSRN: http://ssrn.com/abstract=1840668

Herzenstein, M., Andrews, R. L., Dholakia, U. M and Lyandres, E (2008), "The Democratization of Personal Consumer Loans? Determinants of Success in Online Peer-to-Peer Lending Communities," *Discussion paper*, University of Delaware

Holtz-Eakin, D; Joulfarian, D and Rosen, H (1994), "Entrepreneurial Decisions and Liquidity Constraints", *RAND Journal of Economics*, vol. 25, pp. 334 - 347

Honjo, Y (2000), " Business failure of new firms: an empirical analysis using multiplicative hazard model", *Journal of Industrial Economics*, vol. 18, pp. 557-574

Inc. (2000) The Inc. 500: America's fastest-Growing Private Companies. Inc. magazine, October 17, 2000, pp 65

Iyer, R., Luttmer, P and Shue, K. (2009), "Screening in credit markets: Can individual lenders infer borrower creditworthiness in peer to peer lending?", Working Paper

Jimenez, G and Saurina, J (2004), "Collateral, type of lender and relationship banking as determinants of credit risk", *Journal of Banking and Finance,* vol. 28: pp. 2191-2222

Jimenez, G; Salas, V and Saurina, J (2006), "Determinants of collateral", *Journal of Financial Economics*, vol. 81: pp. 255-281

Keasey, K and Watson, R (2000), " Interest rate premia on UK small firm bank borrowing: A research note", *Journal of Business Finance & Accounting*, vol. 27, pp. 247 - 259

Klafft, M (2008), Online Peer-to-Peer Lending: A Lenders' Perspective (July 14, 2008). Proceedings of the International Conference on E-Learning, E-Business, Enterprise Information Systems, & E-Government, EEE 2008, H. R. Arabnia & A. Bahrami, eds., pp. 371-375, CSREA Press, Las Vegas 2008. Available at SSRN: http://ssrn.com/abstract=1352352

Kumar, S. (2007), "Bank of one: Empirical analysis of Peer-to-peer marketplace", *Working Paper*, AMCIS 2007 Proceedings, Paper 305.

Landis, J. Richard and Gary G. Koch (1977), "The Measurement of Observer Agreement for Categorical Data," Biometrics, 33 (1), 159–74

Lambert, D (2010) " Historical Returns in Angel Markets", Last accessed 20 November,2013 http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0CC8QFjAA&url=http%3A%2F%2Fwww.growthink.com%2FHistoricalReturnofAngelInvestingAssetClass.pdf&ei=r-aQUpilL4LV0QXrlIDYCQ&usg=AFQjCNGKJH5vsEdNYnHG97aPfDSvkzH-vA&sig2=ld5L3AeIElxyd_fyqP0Obw&bvm=bv.56988011,d.d2k

Lehner, O (2013), "Crowd funding social ventures: A model and research agenda", Venture Capital, vol. 15, no 3

Lin, M., and Viswanathan, S (2013) "Home Bias in Online Investments: An Empirical Study of an Online Crowd Funding Market", working paper, available on papers.ssrn.com

Machauer, A and Weber, M (2000), " Number of bank relationships: an indicator of competition, borrower quality or just size? " *Working Paper*, University of Mannaheim

Mallick, R and Chakraborty, A (2002), "The Small Business Credit Gap: Some New Evidence," Finance 0209008, EconWPA

Mollick, R (2013) Swept Away by the Crowd? Crowdfunding, Venture Capital, and the Selection of Entrepreneurs, available at SSRN: http://ssrn.com/abstract=2239204 http://dx.doi.org/10.2139/ssrn.2239204

Massolution, (2013), The Crowdfunding Industry Report, last accessed 21 December 2013http://www.compromisoempresarial.com/wp-content/uploads/137356857-Massolution-2013CF-Excerpt-Revised-04182.pdf.

National Credit Union Association (2008),"Credit Union and Bank Rates: reports and raw data, http://www.ncua.gov/DataApps/Pages/CUBNKMain.aspx, data last accesses on 14 April 2012

Parker, S. (2009), "The Economics of Entrepreneurship", *Cambridge University Press*, pp. 233-281.

Parker, S. (2002), "Do banks ration credit to new enterprises, and should governments intervene?" *Scottish Journal of Political Economy*, vol. 49, no. 2

Petersen, M and Rajan, R (1994), "The benefits of lending relationships: Evidence from small business data", *Journal of Finance*, vol. 49, pp. 3–37

Petersen, M and Rajan, A (1995), "The effect of Credit Market Competition on Lending Relationships", *Quarterly Journal of Economics*, vol. 110, no. 2, pp. 407 - 443

Preston, S. (2007), " Angel finance for entrepreneurs: Early stage funding for long term success", Jossey-Bass, *A Wiley Imprint*, *San Francisco, CA*

Pope, D and Sydnor, J (2011), "What "s in a picture? Evidence from Prosper.com", *Journal of Human Resources*, vol. 46, no.1, pp. 53-92

Ravina, E (2008), "Love & Loans: The Effect of Beauty & Personal Characteristics in Credit Markets" *Working Paper*, Columbia GSB

SBA (2012), "Small Business and Micro Business Lending in the United States for Data Years 2008-2011", A Compendium of Research by the Small Business Administration "s Office of Advocacy, http://archive.sba.gov/advo/research/, report last accessed 22 April 2011

SBA (2010)," Small Business and Micro Business Lending in the United States for Data Years 2008-2009,", A Compendium of Research by the Small Business Administration "s Office of Advocacy, http://archive.sba.gov/advo/research/, report last accessed 22 April 2011

Schwienbacher, A. and B. Larralde (2012), "Crowdfunding of Entrepreneurial Ventures," in D. Cumming (ed.), The Oxford Handbook of Entrepreneurial Finance, Oxford, UK: Oxford University Press

Scott, J and Dunkelberg, W (1998), " Bank consolidation and small business lending: a small firm perspective", In Business access to capital and credit (eds J. Blanton, A. Williamson, and S. Rhine) pp. 238 -361. A Federal Reserve System Research Conference

Sharpe, S (1990), "Asymmetric information, bank lending and implicit contracts: a stylised model of customer relationships", *Journal of Finance*, vol.45, pp. 1069 - 1087

Shane, S (2008), "Fool's Gold: The truth behind Angel Investing in America", *Oxford University Press, USA*

Sohl, J (1999), "The early stage equity market in the USA", *Venture Capital*, vol. 11, no.2, pp.101-120

Solon, A (2012), "Government to invest in Start-ups through P2P lending, Report last accessed on 07 June 2012, http://www.wired.co.uk/news/archive/2012-05/31/government-invests-in-peer-to-peer

Sonenshein, S., Herzenstein, M., and Dholakia, U. M. (2011), "How accounts shape lending decisions through fostering perceived trustworthiness", Organizational Behavior & Human Decision Processes, vol.115, no.1, pp. 69-84SBA (2011)," Small Business Lending in the United States, 2009-2010, Small Business Administration "s Office of Advocacy, http://archive.sba.gov/advo/research/, report last accessed 22 April 2011

Steier, L and Greenwood, R (1995), "Venture Capitalist relationships in the deal structuring and post-investment stages of new firm creation", Journal of Management Studies, vol.32, pp.337-357

Stuart, T, Hoang, H and Hybels, R (1999), "Interorganizational endorsements and the performance of entrepreneurial ventures ", *Administrative Science Quarterly*

Storey, D. J (1994), "Understanding the small business sector", Routledge, London

Stiglitz, J and Weiss, A (1981), "Credit rationing in markets with imperfect information", *American Economic Review*, no. 71, pp. 393-410

The Economist (2012), "The new thundering herd, Wanted: small sums of money to finance young companies", (June, 2012), http://www.economist.com/node/21556973, Article last accessed 30 June 2012

Weiß, N, Pelger, K and Horsch, A, (2010), "Mitigating Adverse Selection in P2P Lending – Empirical Evidence from Prosper.com, available at *SSRN*: http://ssrn.com/abstract=1650774

Wette, H (1983), "Collateral in credit rationing markets with imperfect information: note, *American Economic, Review*, vol.73, pp.299 -313

Zacharakis, A; Meyer, G and DeCastro, J (1999), " Different perspectives of new venture failure: a matched explanatory study of venture capitalists and entrepreneurs", *Journal of Small Business Management*, vol.3, pp.1-14